MULTIVARIATE EXPERIMENTAL CLINICAL RESEARCH

VOLUME 2 NUMBER 2 1976

A Journal for Basic Behavioral Research into Personality Dynamics and Clinical Psychology

Published Quarterly by the Psychology Press, from the Department of Psychology, Wichita State University, Wichita, Kansas, U.S.A.

POLICY STATEMENT

The primary aim of *Multivariate Experimental Clinical Research* is to provide a publication outlet for research in the areas covered and indicated currently by the terms personality study, clinical diagnosis and therapy, extending into the learning, social, physiological, applied and developmental aspects of these. Although due representation is given to theoretical articles which may have a methodological basis, the journal is not one of multivariate statistical methods. Although multivariate in outlook, both manipulative and non-manipulative research is accepted. In fact preference is given to dynamic, manipulative and time-sequential studies. Particular encouragement is provided for pioneer experimental attacks on what is designated personality dynamics and motivation, as well as the natural expansion thereof into structured learning theory.

Multivariate Experimental Clinical Research, Volume 2, Number 2, 1976, 63-88

ALIGNMENT OF PERSONALITY SOURCE TRAIT FACTORS FROM QUESTIONNAIRES AND OBSERVER RATINGS: THE THEORY OF INSTRUMENT-FREE PATTERNS

Raymond B. Cattell, University of Hawaii,
George Pierson, Southern Oregon University, and
Carl Finkbeiner, University of Illinois

ABSTRACT

Testing the hypothesis of "indifference of indicator," i.e., that instrument-free source traits exist as factors when different media are jointly factored, the investigators scored young adults on 32 observer-rating and 32 questionnaire trait markers, known from previous work as markers of personality factors in each medium. The Q-data consisted of A and B forms of the 16 P.F. The L-data was gathered according to six principles considered essential to sound rating, e.g., no fewer than eight peer raters for each subject.

The reliabilities (equivalence coefficient) were about equally good for the two media. Factor analysis revealed 23 factors — the number now commonly recognized in the normal personality sphere. At the unique resolution position of maximum simple structure, the congruence coefficients of (a) L variable loading patterns with previous L-factoring, (b) Q markers with previous Q-factoring, and (c) both kinds of markers with the only available previous Q-L factoring, were highly significant, showing that each of at least the largest sixteen primary source traits has simultaneous, matched expression extending across Q and L media. The possibility that some of the remaining and unidentified seven factors are pure instrument factors is discussed.

THE THEORY OF INSTRUMENT-FREE SOURCE TRAITS BEING TESTED

For many years the senior author has maintained (Cattell, 1946, 1957) what was originally called the "principle of indifference of observation medium," arguing that essentially the same human personality factorial structures will appear in analyses of L-, Q-, and T-data (Life rating, Questionnaire and Objective — Laboratory — Test Variables). That the facts support the principle has, however, frequently been questioned, e.g., by Becker (1960), Howarth (1972), and Fiske (1971), while Passini and Norman (1966) have suggested that even if L- and Q-patterns align it may be because of the structure of words.

Since the issue of the nature and reality of personality trait structures is a vital one for many divisions of psychology — notably clinical, personality, social and industrial — it is an amazing reflection on psychologists' perspectives in research that thirty years have been allowed to elapse since the last methodologically, tolerably adequate experiment on this matter was conducted.

Before proceeding it should be made clear that the "indifference" principle has not remained in the meantime as a bare, non-operational generalization, as it first was, but has developed considerably. Principally it has advanced by the conception that beyond the common factors the different media will have

different *instrument factors* (Cattell, 1973). Defining these, it asserts that aside from these instrument factors common, invariant patterns will appear which may be called *instrument-transcending* or *medium-transcending trait patterns,* corresponding to real underlying behavior structures. This modern form, briefly designated the *theory of instrument-free source traits* (formerly "indifference of indicator"), has developed as part of a still wider rationale of *the relativity of immediate phenomenally perceived factor patterns to the nature of the observers,* called *perturbation theory* (Cattell and Digman, 1964). Therein the perturbation of the true patterns and values by additions due to the nature of observers, modes of observation, and situational settings, are, as in astronomy, systematically conceptually analyzed and given operational meanings. The domain of perturbation theory which most concerns the present research is the taxonomy of instrument factors (Cattell, 1968, 1973). Parenthetically, Campbell and Fiske (1959) have offered probably the only other systematic treatment of these instrument phenomena, in the *multi-trait multi-method* approach. This is not inconsistent with the present conceptualization but it lacks the precision of perturbation theory and instrument factors and has the more limited aim of getting a somewhat improved freedom from instrument effects in psychometric measures of the true trait, without actually locating the instrument factors.

More specifically, Cattell's theory of *instrument-transcending factors,* taken in the personality modality, is that the primary factors on the personality sphere basis in the Q-medium which now run A through Q_6 (15 in the 16 P.F. and 7 in the 16 P.F. Supplement) match similar primaries in the L-medium (life-setting, criterion behavior, either measured or rated by observers). It concludes furthermore, that alignment with the third medium (objective test, T-data) exists, but with a special translation, namely that the second order factors among the L- and Q-primaries match first order factors in T-data. An hypothesis in terms of a rationale for this displacement has already been given (Hundleby, Pawlik and Cattell, 1965).

Matching, in any setting, can be conceived in two ways: as a perceived similarity of psychological meaning and as a statistical demonstration of adequate correlation of the two factors. The first claim for the "indifference of medium" principle was based on (a) the discovery that the number of factors in L- and Q-analyses appear to be much the same (12 to 16), and (b) on their *psychological* meanings permitting them to be ordered in a one to one relationship (Cattell, 1946, 1969). In this initial claim it was recognized and emphasized, however, that matching by "psychological meaning" even as done by careful matching by psychologists, is wide open to abuses of subjectivity. Consequently, not so much weight would have been put on the above conclusion had not the *second order factor patterns* in L-primaries, as identified by the above approach, proved, statistically, almost identical in loading with the second orders found among the hypothesized corresponding Q-primaries (Cattell, 1957). From this earlier stage of method we move today to objective matching procedures, as below.

The three operationally possible bases for objective matching have been set out recently in several places (Buss, 1972; Cattell, 1965, 1969; Kaiser, Hunka and Bianchini, 1971; Nesselroade and Reese, 1973; Schneewind and Cattell, 1970). Although matching factor *patterns* by congruence and *s* coefficients is

reasonably good (Cattell, Balcar, Horn, and Nesselroade, 1969), a somewhat sharper test is provided by the *factoring together* of variables from both or all three media. This was carried out first by Cattell and Saunders (1955), for all three media, by Cattell for Q- and L-data (1968), and by Cattell for Q- and T-data (1955). The first gave a null result; the second, which recognized instrument factors, showed several good L-Q alignments, notable for ego strength, C, excitability, D, surgency, F, and parmia, H; and the third showed very clearly some Q-T matches, notably in revealing the familiar loadings of exvia and anxiety as second order Q-data patterns on the same factors as were marked by the T-data subtests. Other investigators — no one of their published studies using sufficiently objective factor analytic procedures and objective matching tests — have concluded that they could find no alignment or have missed some of the stated primaries themselves in the separate Q- or L-media. Among the former, that by Cattell and Saunders must be included; though by way of apology it must be said that the pioneer study was early in the game (1955), and the null result is now readily seen to be due to the lack of sophisticated methods and really good markers at that time.

Indeed, as to negative results generally it must be pointed out that in this field of integration of factor studies from different media, there are so many steps between the data and the final conclusions, each one of which can and does go wrong in well-known published studies, as Vaughan (1974) has documented, that convergence can appear only through exceptionally skilled and careful procedures. In fact, it can be shown that, due principally to shoddy rotation practices, statistically significant matching is often not achieved even where the very same factors are known to exist, through having constructed plasmodes with built-in matches. In the perspective of a truly evaluative review one must conclude that this handful of crude studies with negative results clashes with recent findings no more than the inability of the professors of Padua to see Jupiter's moons through Galileo's telescope disproves modern research on those moons.

NECESSARY CONSIDERATIONS IN PLANNING AN L-Q ALIGNMENT EXPERIMENT

If one asks where the greatest obfuscation basically originates in past L-Q experiments, the answer (apart from their taking short cuts in factor analysis) is surely in the general weakness of the behavior ratings by which they sought to represent L-data. In spite of the required conditions for sound rating being well-known, and set out several times (Cattell, 1946, 1957, 1973), the survey for the present research has failed to find any personality rating factoring between the studies of 1945 and the present (with the exception of Tupes and Norman) which observed even a reasonable majority of these conditions. The desirable conditions are that the rating be done:

(1) Not on single words, but on *definitions with concrete behavioral illustrations*, to ensure that terms mean the same for all judges and are anchored in behavioral observations.

(2) Not by persons in special role, e.g., authority, relations; but by peers.

(3) Not by as few as two persons (and certainly not by one only), but by at least eight, and yet not by more than can know the man well. Reliabilities are then to

be calculated between averages from two pools of observers, and in such conditions they are usually quite good.

(4) A situational contact with the observed person that is as *broad* as possible, i.e., not restricted to one role segment of his daily life, e.g., behavior on the parade ground.

(5) Utilization of a procedure of ranking all persons on one trait at a time, not all traits on one person at a time.

(6) Removal of inter-rater differences of *means* and, if possible, of *standard deviations* by equalizing these across raters or rated sub-groups.

The reliabilities reached by these methods in the early studies by the present senior author, on young men in fraternities (1947), women in sororities (1948), and tank corps men (1945) in World War II were surprisingly impressive and strongly supportive of the methods above. But, as indicated initially, those methods have scarcely been followed since. Rehabilities for the women (1948) ranged across various traits from .74 to .90 with a median at .855. (A one-hour intelligence test in the same group had a reliability of .82.) Except for the studies of Tupes and Cristal (1961) and Norman (1961) such reliabilities seem not to have been reached since in ratings by poorer methods, but nevertheless used as the basis in factor analytic conclusions. The rather more than 12 primary factors now recognized in L-data rest mainly on these three adult studies, on the studies of Tupes and Norman, and on similarly based studies at the child level by Digman (1963), Peterson and Cattell (1958), Coan and Cattell (1957), and others.

Even when the rating side is well designed and carefully done, however, difficulties of factor resolution in the combined L-Q domain can still arise from other sources. They arose, for example, in the earliest matching studies from the primitive questionnaire scales marking the factors also having prominent loadings on other factors than those aimed at. It has taken more than fifteen years of *progressive rectification* (Cattell, 1973) to reach relatively factor-pure questionnaire scales in a dimensionality of, say, 14 or more distinct source traits. Difficulties from such composite loadings of early CPQ scales arose, for example, in the otherwise basically well-designed L-Q study by Schaie (1962). This study showed essentially a match of L- and Q-primaries across the whole spectrum, but nevertheless with quite low and uneven loadings. To critics who deny that promising leads to cross identification can be inferred from low correlations and to Becker's criticism (1960) that correlations of only around .3 were found between the L- and Q- markers for the same factor, Cattell has pointed out the need to consider the correlation matrix as a whole. For example, an r of only .36 between two variables supposed to mark the same factor corresponds to a loading of .6 by each of the two variables on that same factor. And a loading of .6 is about as much as is commonly found either with teacher ratings or with short Q-scales when we are dealing with variables as markers even in one and the same medium. The real requirement in planning and interpreting an experiment across media, with the average available markers, loading .6 or a little more, is not that really *high* correlations be reached between L- and Q-markers for the same factor, but that the L- and Q-markers correlate consistently in relation to known factor loadings and reliabilities. In fact, variables hypothesized to have significant loading on one factor all fall with moderate loadings on that factor and essentially nowhere else. Here one notes that an important part of the requirement is that these

cross media matching variables should be in the hyperplane of, i.e., have no loadings on, other L— and Q-factors *except* instrument factors.

As to planning to recognize instrument factors themselves, by inserting markers for them, one can do little at this stage because their natures in the L- and Q-domains are as yet not well known. They must be considered something to be discovered in such studies as the present by simple structure rotation and tested as hypotheses in later experiments. Almost certainly — if *trait view theory* (Cattell, 1968) is correct — (a) there will not be *one* instrument factor for each medium, L and Q, simply running evenly across the whole rating or questionnaire domain. However, on the other hand, we would not expect one to *transcend* a domain, and (b) it is a mistake to regard L-data (or, at least, *rating* L-data as distinct from behavioral time-sampling L-data) as "the criterion," i.e., as entirely free of instrument factors, and Q-data as "the test." For perturbation theory tells us that L-data probably has larger instrument factors than Q-data, some tied up with the rather specialized settings of real life behavior and observation.

The more specific, more self-evidently necessary conditions in the planning of this experiment, e.g., having a minimum of two markers in each medium for the statistical requirement of marking each factor, need not be explained, being often previously stated in factor analytic work of this kind, and are merely recorded below in the experiment.

THE EXPERIMENT: SUBJECTS, VARIABLES, DATA GATHERING

Since factor analyses tend to lack stability below about 130 subjects, we aimed at a sample of 150, and since we wished generalizability across the sexes, we took men and women in about equal numbers, ending with a sample of 76 young men and 56 young women. The subjects were mainly undergraduates from a small college in a small town in Oregon, where opportunities for subjects to study one another were good, and to these were added 25 undergraduates in psychology at the University of Illinois belonging to the same dormitories.

The rating variables were taken from the basic set gradually accumulated over years to represent the personality-sphere. Forty-two such variables are set out with definitions in Appendix 2 of Cattell (1957, p. 813). This list has been used in personality-sphere investigations quite steadily since, e.g., by Norman, Tupes, Digman, and others. As a standard reference set of definitions covering most personality behavior and offering comparability with other extensive researches, it seemed well to use it with minimal change. The only modifications of the list here were (a) to drop variables that had not consistently shown steady loadings at a high level as markers on the main rating factors A through O, and (b) to add rating variables which correspond to the meanings now given to the Q factors (Q_1 through Q_6) which, as their Q designation indicates, have not previously been caught in rating studies. We planned in fact here for the first time to test our hypothesis that these "extra" Q factors (Q_1 through Q_6, Cattell, 1973) had previously only been missed in pure L-data analyses through accidentally insufficient coverage of personality-sphere variables. The list used is given in Table 1. Each variable is presented with a definition and illustration of two opposite poles. Incidentally, this does not imply "positive" behavior at both poles, and "Intelligent-vs-Stupid" as a bipolar rating still involves only the

conception of something called intelligence which goes in one direction from a zero to a high value. The practice is followed because it has been found that absence of definition of the opposite pole leaves real inadequacy in the logical precision of the vector concerned. It will be noticed that the *second* member of each pair of markers is arranged to be rated *high* by the raters from the *opposite* pole to that which would correlate positively with the first variable in the pair. This is part of the design to eliminate "halo." They were presented to the raters in irregular order, not in the pairs as arranged meaningfully for the reader in Table 1.

Table 1

List of Bipolar Rating Variables*

Factor for which it is a marker	Number in factor Table 3	
A	1. (28)	Good Natured, Easy Going. Does not object when people use his property, time or energy; generous and warm hearted in his general attitudes; gives people the "benefit of the doubt" when their motives are in question. vs. Critical, Exacting. Likes people to be precise and dependable; does not tolerate human failings; is skeptical about people and promises; requires proof of suggestions put to him.
A	2. (13)	Cool, Reserved. Tends to be indifferent to personalities and to ignore people, "standoffish," when personal matters are discussed; cold and aloof toward others; does not mix when in a crowd. vs. Interested in Others, Warm. Attentive to people; makes friends easily and quickly; listens to and participates in interests, problems and concerns of others.
B	5. (17 in new list)	Can Think Abstractly. vs. Has Difficulty with Ideas Concretely Set Out.
B	6. (48 in new list)	Slow to Learn vs. Learns Fast
C	9. (49)	Mature vs. Changeable, Erratic
C	10. (2)	Emotional vs. Stable, Deliberate
E	13. (11)	Assertive vs. Submissive

Table 1 Continued

E	14. (9)	Self-Effacing	vs.	Egotistical
F	17. (26)	Merry	vs.	Sober, Solemn
F	18. (22)	Prudent, Careful	vs.	Happy-Go-Lucky
G	21. (3)	Conscientious	vs.	Unconscientious
G	22. (7)	Quitting	vs.	Persevering, Responsible
H	25. (41)	Frank, Friendly	vs.	Secretive, Shy
H	26. (20)	Nervous, Timid	vs.	Bold, Adventurous
I	29. (8)	Tender	vs.	Tough, Hard
I	30. (18)	Ready for Whatever Comes	vs.	Demanding, Dependent
L	33. (5)	Prone to Jealousy	vs.	Not Jealous
L	34. (25)	Trusting	vs.	Suspicious
M	37. (38)	Imaginative, Fanciful	vs.	Practical
M	38. (4)	Conventional	vs.	Bohemian
N	41. (51)	Astute, Artful	vs.	Spontaneous, Natural
N	42. (32)	Naive, Genuine	vs.	Socially Skillful, Polished
O	45. (54)	Worrying and Depressed	vs.	Self-Confident
O	46. (53)	Self-Assured, Complacent	vs.	Guilty and Easily Ashamed
Q_1	49. (55)	Liberal, Experimenting	vs.	Accepting Tradition
Q_1	50. (56)	Conservative, Accepts What Is	vs.	Radical, Questioning Ideas
Q_2	53. (57)	Resourceful, Takes Initiative	vs.	Dependent on Group, Asks What Others Do
Q_2	54. (17)	Socially Dependent	vs.	Self-Contained, Makes Up His Own Mind
Q_3	57. (29)	Insistently Orderly, Plans Ahead	vs.	Does Not Think of Consequences, Scatterbrained or Disorganized
Q_3	58. (34)	Casual, Careless of Public Opinion	vs.	Mannerly, Shows Regard for his Social Image

Table 1 Continued

Q_4	61. (24)	Tense, High Strung	vs.	Relaxed, Seems Under No Pressure
Q_4	62. (58)	Phlegmatic	vs.	Overwrought, Irritable, Loses Temper Easily

*These 32 variables were given in irregular order to raters, not paired as here. Each member of the pair was rated high in an opposite direction from the other, the italicized being high.

The numbers in parentheses are those given these variables in the standard table of personality sphere variables (pp. 813-817, Cattell, 1957) to which those of Norman's variables not well represented (half a dozen) have been added in continuation numbers. The letters on the left are, of course, the usual alphabetic primary factor identifications.

The more concrete behaviors for a trait, as given the raters for illustration and definition, are shown here, for economy, only with the two first variables. The complete rating list of concrete behaviors, along with further details of procedure, may be obtained from Drs. Pierson, Brim, Robertson or Willes at Southern Oregon College, where most of the rating was done.

The procedure in the rating was intensive. The subjects were in groups of 10 (occasionally down to 9) and each group met 36 times (a total of about 30 hours) over the semester in situations akin to therapy groups, in which each subject would discuss his problems and his personality, and in which analytical attention was also given to the meaning of the variables as such and the associated behaviors (Table 1) with illustrations by which each would be assessed. Anyone not wishing to participate was allowed to drop out. In the actual rating, done at the end of these sessions, each rater was anonymous (given a number only) and the traits were taken one at a time (about 5 to a session) and, as stated, in random order relative to Table 1. There remained at the end of the semester in every case at least 9 persons to rate each person in each group and for uniformity only 9 raters were used from any group. The procedure was thus simply of one *ranking* 8 people, on each trait, carried out in individual isolation by 9 people in each group (the 9th person in the group not rating himself but only the other 8). Then, assuming a normal binomial distribution — 1:3:3:1 — over the 8 subjects, each rater's rank order was converted to a point score of 1, 2, 3 or 4 on each of the 32 traits. To get each individual's final score these point scores (1) were added for each of two groups of 4 randomly divided from the 8; (2) the rating reliability was calculated between these two pools (Table 2 below), and (3) the two were finally thrown into a single score.

Corresponding to the representation of rating factors by these 32 scores, the hypothesized rating factors were also each marked by two variables. These were the A and B form scores for each of the 16 P.F. factors, A through Q_4. This

involved some two hours of testing, given close to the end of the sessions for each subject, who was assured of anonymity, though allowed after the experiment to see his own score.

Table 2

Internal Consistency of Rating and Questionnaire Variables

(a) Ratings. Homogeneity or "Generalized Reliability."

Factor Marked	No. of Variable		I	II	No. of Variable		I	II
A	1	Good natured	.58	.59	17	Cool reserved	.64	.65
B	2	Thinks well abstractly	.78	.78	18	Slow to learn	.74	.74
C	3	Mature	.66	.67	19	Emotional	.66	.67
E	4	Assertive	.88	.89	20	Self Effacing	.67	.68
F	5	Merry	.81	.81	21	Prudent, careful	.66	.67
G	6	Conscientious	.55	.56	22	Quitting	.71	.72
H	7	Frank, friendly	.85	.85	23	Nervous, shy	.80	.80
I	8	Tender	.73	.74	24	Ready and tough	.64	.65
L	9	Prone to jealousy	.62	.63	25	Trusting	.63	.64
M	10	Imaginative, fanciful	.61	.62	26	Conventional	.68	.69
N	11	Astute, artful	.65	.66	27	Naive	.69	.70
O	12	Worrying, depressed	.72	.73	28	Self assured	.61	.62
Q_1	13	Liberal, experimenting	.73	.74	29	Conservative	.73	.74
Q_2	14	Resourceful	.78	.78	30	Socially dependent	.64	.65
Q_3	15	Insistently ordered	.70	.71	31	Casual	.76	.76
Q_4	16	Tense, high strung	.54	.55	32	Phlegmatic	.66	.67

As mentioned in the text, column I is Spearman-Brown and Kristof maximum likelihood, while II is the latter converted to an unbiassed estimate.

Mean general reliability for ratings = .70

(b) Questionnaire Scales. Reliabilities calculated as dependability coefficients.

		Form A	Form B
A	Affectia	.81	.75
B	Intelligence	.58	.54
C	Ego Strength	.78	.74
E	Dominance	.80	.80
F	Surgency	.79	.81
G	Super Ego Strength	.81	.77
H	Parmia	.83	.89
I	Premsia	.77	.79
L	Protension	.75	.77
M	Autia	.70	.70
N	Shrewdness	.61	.60

Table 2 Continued

O	Guilt Proneness	.79	.81
Q_1	Radicalism	.73	.70
Q_2	Self Sufficiency	.73	.75
Q_3	Self Sentiment	.62	.62
Q_4	Ergic Tension	.81	.87

Mean dependability reliability for questionnaire scales = .74

Most calculations allowing for "error" in the L and Q measures would require the dependability measure, as in 2(b). Since it was not possible to get the same subjects re-rated by the same judges it has seemed best here to take in 2(a) the same rated by different judges and consider the error to reside in the difference in people's perception of the same defined variable.

A feature of the treatment of scores here, over and above the general precautions above, which may have helped to account for the more clean-cut results, is the introduction of a correction for inter-group variation. It will be noted that the above rating procedure results in each sub-group of 8 subjects having the same mean though not the same sigma. The argument for this is that though there are probably small real differences of means between sub-groups the actual raw score mean differences usually found are fairly large and mainly due to differences in "leniency" of raters. This contamination of inter-subject real differences with inter-rater spurious, intrusive variance would naturally introduce a false pattern of variance-covariance in the trait scores. Moreover, there are plenty of reasons for doubting that this rater covariance is of the same structure as the true subject variance-covariance, i.e., it will be seriously distortive. It pays, therefore, to get rid of it even at the cost of some slight loss of the true variance-covariance which would exist between subject groups. For according to theory the factor structure of the between group, real variance-covariance should, with random groups, be from the same sources and the same nature as the within (inter-individual) variance-covariance. In getting rid of it, along with the spurious part, therefore, we merely reduce somewhat the range and therefore the size of the absolute loadings, not their true *patterns* on the factors.

However, if the rating data is treated in this way, it becomes essential to treat the questionnaire data in the same way, so that there is no unmatched variance-covariance in one medium. It is true that in this Q-data case we are unfortunately getting rid of some real variance-covariance, but this slight loss is worthwhile in terms of the clearness gained in the overall design. Incidentally, there is also an argument for bringing the sub-groups to the same *sigmas*, as well as *means*, in ratings, before throwing all sub-groups' scores together, on the grounds that groups having more reliable judges — not merely more extreme people — will have larger raw score sigmas. However, with present methodological knowledge we decided not to experiment with this refinement, though our hunch is that the question deserves statistical examination.

On a qualitative level of observation we would note that the use of *repeated group discussions on personality description in general and in regard to particular behavior,* increased interest and accuracy considerably. Both the

questionnaire and the rating operations, moreover, were more conscientiously carried out than we have sometimes seen, frankness being engendered by the anonymity and keenness by the sense of contributing to an extensive group research.

THE ANALYSIS

As to the consistency of the data we have calculated reliabilities — strictly *equivalence coefficients* as a form of *homogeneity* (Cattell, 1973) — for the ratings in Table 2(a) and as *dependability coefficients* for the questionnaire scales in 2(b). The latter require no explanation beyond pointing out that since we did not repeat each scale here the values are taken from the 16 P.F. Handbook for a student group closely equivalent to that used here. The former, however, because of the rarity of rating evidence of this kind, were calculated, for general psychological interest, in several ways: (a) as a Spearman-Brown extension to the whole (8 raters) of the r between the two parts (of 4 raters each), (b) as Lord and Novick's (1968) "generic reliability" (Cronbach's "index of generalizability"), (c) Kristof's maximum likelihood estimate (1973), (d) Kristof's estimate for the lower bound at the 95% one-tailed confidence interval, and (e) Kristof's conversion of the maximum likelihood to an unbiased estimate. The order of equivalence of the two rating sub-groups was practically the same for all, and, in the case of (a) and (c) identical to two places. Consequently, and because (a) is most familiar, we have set out only (a) and (e) estimates in Table 2(a).

It will be observed that if we consider the 16 P.F. scales as measuring literally what they measure, i.e., varying as to reliability but not validity (so that error is error in repeated measurement, i.e., is evaluated by the dependability coefficient) and if we define the rating error as the extent of human misunderstanding and disagreement concerning a defined trait (inverse of equivalence), then ratings, by our method, and questionnaire measures are *essentially* at the same level of test consistency. However, the ratings are *slightly* lower than the questionnaire scales and, incidentally, these 8-judge ratings are slightly lower than the 10-judge pooled ratings we obtained in the three studies thirty years ago (Cattell, 1945, 1947, 1948). Any inferences from this for factor loading variations are made under discussion of loadings below. In connection with the theorem that in Spearman-Brown calculations number of judges and number of items act similarly, we had here 8 people on one side (L-data) and 10-13 items on the other (Q-data). It is interesting to note that the S-B "correction for length" brings the consistency coefficients for L- and Q-data virtually to identity. It is also noteworthy, but not for further exploration here, that certain factors tend across media to have high consistency in their marker variables, e.g., factors E, H, O and others, e.g., factors N and Q_3 to be low, as if more subtle in nature. As to other properties at the sheer measurement level the raw score range of L-scores in the given units was about 2½ times that of Q-data. This might contribute to Q-intercorrelations in general (in absence of correction of r's for grouping) being found somewhat lower. The sigmas were very even across all L- and across all Q-variables, except for intelligence (B in the 16 P.F. test; Q-data) which was lower, probably due to using a small-range student group.

The 32 L- and 32 Q-variables were product-moment intercorrelated in a 64 x 64 matrix, where it was reliably observable that the correlation among the L-

variables averaged (arithmetically, sign ignored) definitely higher than among the Q-variables. Since the subsequent analyses shows no more factors among them, and no higher loading on the factors to be marked, it seems unlikely that the above difference in range is responsible for higher average correlation. The result is clearly due to the ratings being less pure measures than the Q-scales in respect to the factors they are entered to mark. Inspection of Table 3 indeed shows that ratings manifest more scattered low loadings on other factors, besides those they should ideally alone mark, as L-variables. This contrast is not surprising when one reflects that the L-variables were originally chosen to stake out the full dimensionality of the personality sphere. Although a minority turned out to be tolerable markers for a special factor they were never pure, whereas, of course, the 16 P.F. scales were shaped from the beginning, by progressive rectification (Cattell, 1973), each as a pure representative as far as possible, of one factor.

The examination for number of factors was made by the scree test, shown in Figure 1. The truly sharp scree ends at 23 factors, and if the first 23 factors were rotated to equal variance they would also satisfy alpha coefficient significance, i.e., be all above the K-G cut-off of 1.0. It is noteworthy that this number 23 is the number repeatedly found (Cattell, 1973) in a comprehensive personality sphere representation in *normal behavior* Q-data (Cattell & Delhees, 1973; Cattell & De Voogd, 1974) (actually the same is found when the roster of 32 or more 16 P.F. scales instead of more numerous items are used, since although they are chosen to mark the clearest 16 they extend to a dimensionality in various studies of 20 to 23 [Kulhavy & Cattell, in press]).

The three L-data studies in the literature using almost equally comprehensive variables converged, on the other hand, upon about 12 factors (Cattell, 1945, 1947, 1948) and several L-studies *without* a check on factor number have stopped short even of 12. At that time there was no really reliable way of assessing number of factors and it was fashionable by "looking at residuals" to stop at half a dozen or even 3 factors! The rudiments of what later became the scree were used in the 1945-8 studies, but since the developed scree has since shown high accuracy in deciding the number of factors where the number is definitely known on ulterior grounds (Cattell & Jaspars, 1967; Cattell and Vogelmann, 1977) we decided to go back and check by a refactoring of the 1945-48 rating data, still probably the best available in the field. The result is shown in Figure 2 and by the ordinary use of the scree it would indicate 23 factors. At the very least there would be 19, by an alternative line as shown dotted in Figure 2, though anyone experienced in scree usage would definitely give this second probability. This work on ratings on the women's sorority group of N = 230 therefore agrees very well with the verdict of 23 factors in the present combined L-Q data, though it remains to be seen *here* whether the *rating* data shares loadings on all 23 dimensions in the L-Q joint media.

One could, of course, factor the L-data here alone, but that is another study, and in the joint media we see every reason to take 23 factors. Accordingly the communalities were iterated to fit that definitive number.

Alternative rotational approaches to simple structure by automatic programs were tried with Maxplane and Procrustes, after which repeated shifts were made by Rotoplot. Fewer overall Rotoplot rotations were made than usual,

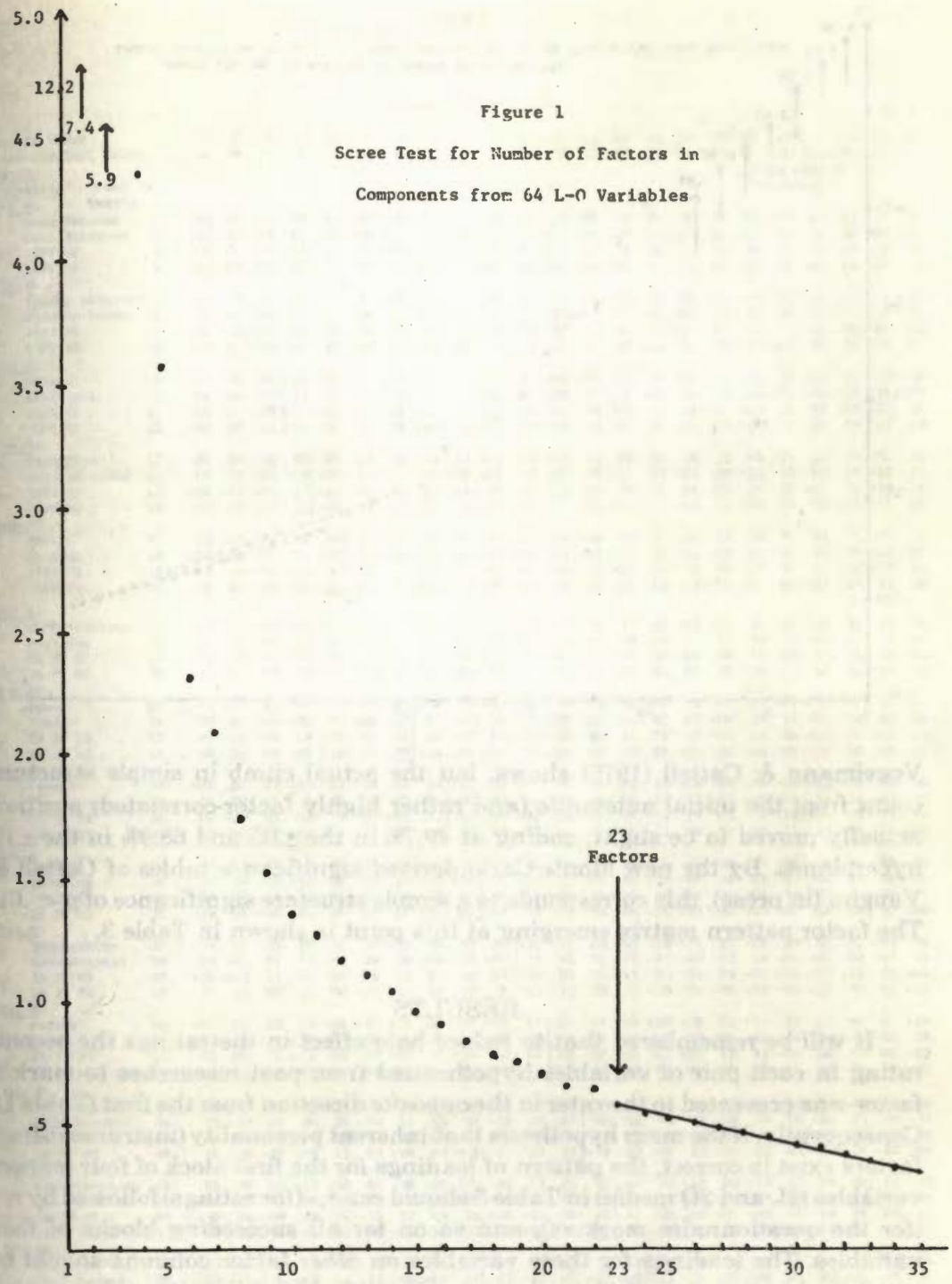
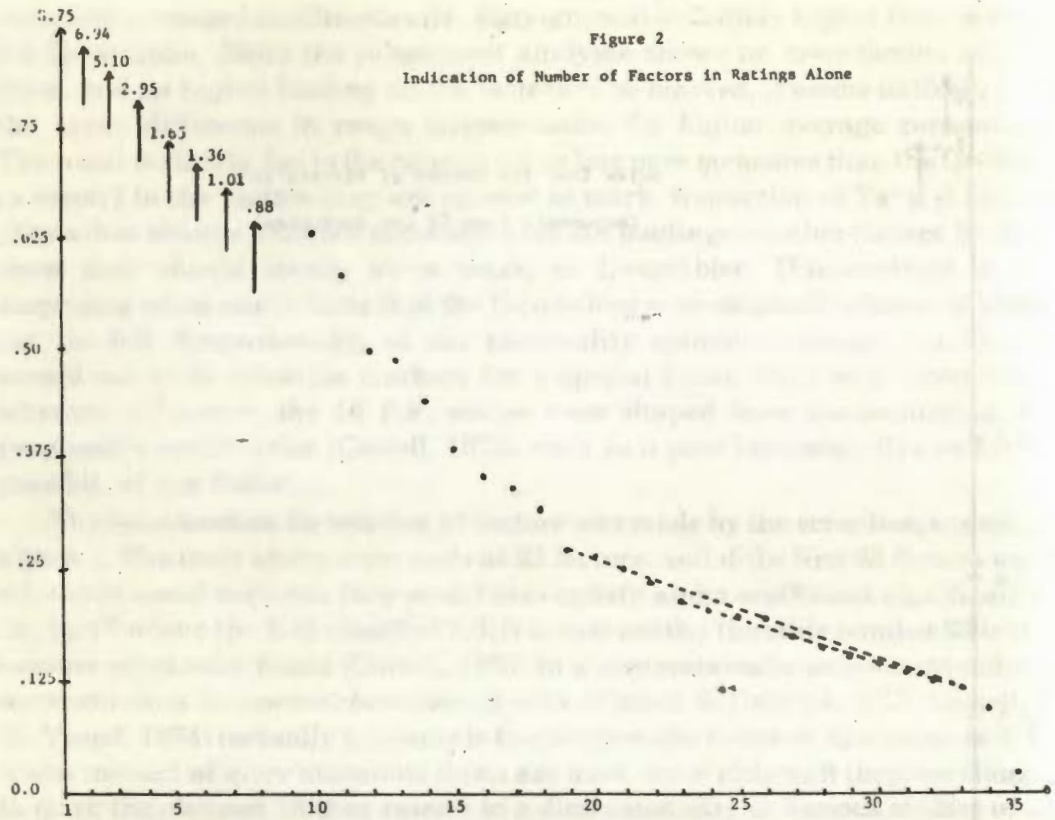


Figure 1

Scree Test for Number of Factors in

Components from 64 L-O Variables

namely nine, before we decided by the plot of "History of Hyperplane Count" that a plateau was reached. As with all rotations with large numbers of factors the main problem in resolution was that comparatively small correlations in the reference vectors led, by inversion, to large correlations among the factors. The straight pursuit of maximum simple structure, as usual, proved the surest way to reduce these angles and they finished at very satisfactory levels as Table in



Vogelmann & Cattell (1976) shows, but the actual climb in simple structure count from the initial automatic (and rather highly factor-correlated) position actually proved to be slight, ending at 49.7% in the ±.05 and 68.9% in the ±.10 hyperplanes. By the new Monte-Carlo derived significance tables of Cattell & Vaughn (in press), this corresponds to a simple structure significance of $p < .01$. The factor pattern matrix emerging at this point is shown in Table 3.

RESULTS

It will be remembered that to reduce halo effect in the ratings the second rating in each pair of variables hypothesized from past researches to mark a factor was presented to the rater in the *opposite* direction from the first (Table 1). Consequently, if the main hypothesis that inherent personality (instrument-free) factors exist is correct, the pattern of loadings for the first block of four marker variables (2L and 2Q media) in Table 3 should run +,- (for ratings) followed by +,+ (for the questionnaire markers), and so on for all succeeding blocks of four variables. The loadings for these variables on *other* factor columns should be small, and in the given factor column, as identified by this pattern of four, any remaining loadings should ideally be zero (i.e., actually, insignificant), as in any good hyperplane.

Inspection of Table 3 shows (see boxes) that this hypothesized structure is essentially followed for some 16 factors (arranged here in the usual alphabetical order) (Cattell, Eber and Tatsuoka, 1970). However, there are scattered departures from this pattern. First, the markers sometimes have loadings also on other factors: a gross instance occurs in the first two rating variables, good natured

TABLE 3

FACTOR PATTERN OF JOINT L & Q FACTOR ANALYSIS AT UNIQUE SIMPLE STRUCTURE RESOLUTION

FACTOR SOURCE TRAITS

Matrix Order		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Psychological Ident.		A	B	C	E	F	G	H	I	L	M	N	O	Q1	Q2	Q3	Q4	Error	Q?	Fake Good.1 [or D(-)]	Fake Good2	X-?	Q?	Q?
Marker Variables*	Order in Matrix																							
For A:																								
L Good Natured	1	20	02	02	-29	52	39	08	-04	-07	08	-21	-20	-06	04	01	-34	03	-02	04	02	-11	04	13
Cool Reserved	2	-10	16	09	01	-53	-44	-17	01	19	02	05	-02	-05	-28	-24	-16	07	06	13	13	07	-07	-11
Q 16PF A1	3	55	01	37	-10	11	08	-06	15	23	-11	02	-24	04	08	12	22	03	34	36	14	08	11	20
16PF A2	4	61	-03	05	09	-14	00	14	04	-06	-07	14	-06	04	-02	14	11	01	23	32	03	20	-12	05
For B:																								
L Thinks abstract	5	13	38	28	26	07	01	-04	-08	-14	04	32	14	-06	28	08	02	-14	00	-06	-12	26	-12	-47
Slow to learn	6	-08	-44	-08	-12	-15	-06	05	-01	-07	-06	-42	04	-06	-40	-21	-02	05	01	04	12	-21	11	38
Q 16PF B1	7	-06	48	-01	10	-06	06	14	03	02	04	18	-20	-11	04	-09	-02	-04	37	07	03	-31	01	21
16PF B2	8	04	62	-03	-01	10	-03	-06	00	-35	-16	-11	09	19	10	-11	00	35	-09	16	-01	22	38	04
For C:																								
L Mature	9	14	06	14	08	08	26	-08	-09	-07	-16	07	-05	07	28	22	-45	-11	-12	10	-14	19	06	-08
Emotional	10	16	04	-16	02	02	17	17	23	19	-06	-27	-08	02	-18	-07	41	09	-02	-32	-14	-15	15	13
Q 16PF C1	11	-14	17	36	03	-03	-08	16	08	-23	05	06	-19	02	-03	17	09	01	-10	03	22	-19	-02	18
16PF C2	12	-06	07	44	01	-05	-02	05	-06	-21	-01	13	-19	10	01	02	-01	18	-10	17	06	19	35	20
For E:																								
L Assertive	13	01	06	06	55	08	21	18	-04	12	-03	28	-02	21	32	04	18	13	04	-04	-15	10	05	10
Self effacing	14	05	02	-07	-69	-04	13	-14	-02	-29	-02	-33	-03	03	-14	03	-05	08	-12	31	-08	-15	-12	16
Q 16PF E1	15	-04	-03	-04	23	-16	-01	65	03	24	07	-12	-12	09	08	-14	04	05	17	18	-02	02	-50	-14
16PF E2	16	-17	02	22	34	-04	18	25	-10	15	-05	-22	06	14	-12	-27	-02	-27	-02	28	09	-01	-21	05
For F:																								
L Merry	17	11	01	04	18	66	12	13	-09	01	07	-01	04	-06	08	05	-08	12	04	01	01	02	10	41
Prudent	18	-30	06	07	-01	-52	13	-16	09	-05	-08	-01	11	-07	-11	26	11	-07	01	13	00	00	-11	04
Q 16PF F1	19	20	04	-01	-12	37	05	30	13	-05	-03	-05	-01	12	-10	03	08	01	-07	77	12	02	-13	21
16PF F2	20	01	02	20	-31	61	-01	22	-05	12	-03	05	07	08	-01	10	03	-01	-01	60	-01	20	12	10
For G:																								
L Conscientious	21	07	14	05	-16	-16	36	-03	10	-07	-15	-08	-02	04	22	36	-23	-01	14	01	-20	11	05	25
Quitting	22	-02	-03	-14	-12	04	-31	-07	-02	-01	08	-09	10	-09	-40	-54	01	10	-09	13	-01	-04	13	22
Q 16 PF G1	23	00	26	-78	01	03	49	02	01	-06	-11	17	-11	08	-10	18	-01	08	11	08	22	43	12	04
16 PF G2	24	11	-08	-09	07	-12	66	-16	06	-02	-01	01	01	-12	-18	03	01	26	-08	12	65	06	-11	-18
For H:																								
L Frank	25	17	-01	03	41	24	44	28	-08	10	-01	08	04	18	16	-01	14	05	01	-02	-13	-01	06	11
Timid	26	-08	-04	-02	-35	-33	-06	-23	02	-02	01	-14	-05	-21	-36	05	01	-04	06	13	-07	01	07	-04
Q 16 PF H1	27	11	-01	04	09	-10	12	84	04	07	00	-13	-20	-04	11	01	08	-01	02	34	-02	26	-04	-04
16 PF H2	28	00	00	14	06	-18	07	89	-04	00	-03	03	01	02	12	08	-07	00	03	10	08	10	08	12
For I:																								
L Tender	29	-01	07	02	-53	23	10	-07	24	-17	15	-36	07	04	04	19	-08	06	-06	03	-02	-06	05	26
Practical	30	14	07	08	35	08	22	00	-04	-04	-04	17	-03	06	34	01	-22	-13	-06	19	11	18	-07	-18
Q 16 PF I1	31	06	-19	05	01	15	-01	-09	77	13	14	-02	04	-05	14	-05	19	03	17	28	-06	24	15	04
16 PF I2	32	13	-04	-19	-02	-10	17	13	63	03	09	-05	12	01	10	-12	03	10	01	-23	13	-09	25	-08
For L:																								
L Jealous	33	-10	-05	02	49	08	-01	-09	02	23	-03	21	-14	-10	-29	-12	33	08	20	-08	06	-03	-09	-12
Trusting	34	14	14	05	-12	46	19	23	10	-38	-04	-17	12	01	06	18	-48	08	07	09	04	07	-12	08
Q 16 PF L1	35	-01	12	-12	-12	01	01	-01	-08	61	00	-26	-05	13	06	08	00	07	-01	15	07	14	-54	-01
16 PF L2	36	-10	05	-04	-01	00	07	01	00	42	-21	05	33	-10	04	-18	-17	-34	-03	01	01	-08	00	-09
For M:																								
L Imaginative	37	11	-01	05	00	24	10	-11	05	07	41	-26	14	20	40	-18	-11	03	-08	05	03	05	-10	76
Conventional	38	00	-01	02	13	-02	04	06	08	-08	-02	15	10	-19	-20	80	-03	02	26	-04	-07	-01	-03	15
Q 16 PF M1	39	-06	-03	11	-04	-06	08	15	09	02	65	08	-08	14	-12	00	05	05	-04	-08	-02	-14	02	-44
16 PF M2	40	01	18	00	27	00	-27	-03	46	-05	30	02	12	32	-03	-05	05	-08	-05	-01	-02	-19	-12	-10
For N:																								
L Astute	41	-07	10	05	35	-06	-23	-05	00	09	-14	67	01	02	10	-05	-08	02	-10	14	-02	07	03	-29
Naive	42	-12	-05	-19	-16	-04	01	05	05	-12	-05	-68	00	-12	-44	-16	03	01	00	-06	02	-07	-04	18
Q 16 PF N1	43	07	00	02	-15	00	06	-07	00	-14	00	41	01	-02	10	02	-02	05	88	-09	00	16	02	-03
16 PF N2	44	07	-10	15	01	-14	14	-29	06	09	07	15	00	01	-43	12	01	00	09	-14	10	14	28	-15
For O:																								
L Worrying	45	-08	06	-03	-48	-29	11	-21	-07	06	04	-14	11	-02	-48	-16	05	03	01	06	02	-18	03	08
Self Assured	46	03	02	06	34	28	14	11	02	04	-10	24	-20	08	26	00	-24	06	04	20	05	11	-20	-11
Q 16 PF O1	47	04	00	-02	-05	08	07	02	09	12	02	00	94	12	07	04	03	07	18	-31	05	25	07	02
16 PF O2	48	14	00	-29	05	12	03	-26	03	-05	10	-03	27	-12	07	-09	22	-04	01	-03	-03	-33	-20	-09
For Q1:																								
L Experimenting	49	05	09	02	-11	40	05	00	00	02	09	12	04	33	47	-32	-12	08	-13	15	06	04	01	23
Conservative	50	-06	-06	07	07	-20	06	-12	03	-06	-10	-24	00	-28	-42	39	-12	-03	07	09	-08	-10	08	26
Q 16 PF Q1 1	51	-30	00	09	03	24	-15	17	05	14	-08	-45	-14	42	12	03	-09	-05	-05	31	02	16	03	09
16 PF Q1 2	52	01	-05	04	05	-02	06	-01	-09	-01	04	-02	02	82	-06	05	05	07	13	00	01	06	-06	-10
For Q2:																								
L Resourceful	53	11	16	13	47	-01	16	11	01	05	09	26	02	19	52	04	06	00	04	-12	-08	13	-02	04
Soc. Depend.	54	00	-11	-06	-42	36	12	08	-07	08	02	-07	05	-12	-45	22	10	23	08	07	02	-26	13	13
Q 16 PF Q2 1	55	-64	00	09	00	-15	-05	-14	-01	05	-02	-04	03	06	22	-08	10	01	14	-04	03	-04	-08	13
16 PF Q2 2	56	-40	-10	12	07	08	07	-32	18	05	-14	-16	-15	05	12	-17	02	-06	30	-04	-11	19	-08	-13
For Q3:																								
L Orderly	57	-14	23	07	-10	-07	18	10	12	-05	-21	46	15	-03	17	64	-06	05	00	08	-14	09	03	-19
Casual	58	-18	09	01	02	-03	-09	12	03	-10	-04	-07	06	12	-08	-87	-01	-09	-06	06	-01	-18	-01	-07
Q 16 PF Q3 1	59	-07	-06	08	06	00	05	-13	05	-03	02	-16	-29	14	18	26	14	-34	03	03	69	00	19	21
16 PF Q3 2	60	-23	00	-02	12	-07	16	-05	04	06	-04	18	-40	-07	01	25	06	22	08	-01	15	45	04	-03
For Q4:																								
L Tense	61	04	02	01	01	-01	10	00	10	06	-04	-25	-08	06	-01	-05	77	03	20	-33	05	00	03	-22
Phlegmatic	62	02	01	14	-22	03	10	00	-08	-09	05	-06	-03	-03	10	-02	-68	00	-02	22	09	-03	-06	09
Q 16 PF Q4 1	63	17	12	-13	04	-03	24	-15	-07	-03	03	-12	48	01	05	-05	39	-02	10	27	-06	-13	-25	16
16 PF Q4 2	64	09	00	-15	07	14	01	-14	01	02	08	-11	38	-07	06	-02	26	03	02	02	-18	-11	-28	-10

* Fuller description of the short titles is given in Table 1.
The variables which should be markers for the factors concerned have their loadings marked by half-boxes.

and reserved (-), which also load surgency (F) and super ego (G) in some instances more than the factor, affectia, A, they were primarily set to mark. This, however, is exceptional, as will be seen by looking at factors intelligence (B), surgency (F), super ego (G), parmia (H), shrewdness (N), self-sufficiency (Q_2), etc., where in no instance is the pattern repeated in the first sixteen substantive factors. Correspondingly, proceeding down the columns, the great majority of loadings — about 80% if we take .13 as the significance limit — on anything other than the hypothesized markers are insignificant.

However, we should also take note of a minority of instances where hypothesized loadings are partly astray. Thus the intelligence factor shows a positive loading of .23 on the questionnaire super ego scale A and this same maverick B scale, though loading on super ego G, where it should, has a large negative loading on ego strength, C. On factor E, besides the appropriate loadings on the four markers, there are large loadings on 25, Frank (.42); 29, Demanding (-.53); 33, Jealous (.49); and 45, Self Confident (-.48). (Note word poles have been adjusted to signs, i.e., negative on 45 is Self Confident.) These are all ratings which have normally been found to have loading on Dominance, E (Cattell, 1957), though they load better on factors for which they are chosen as markers. As frequently happens, even with objective tests (Cattell, Horn and Sweney, 1964) the self-sentiment factor Q_3, has some loading effect on super ego, G variables. The remaining departure from cleanest structure in terms of hypotheses is also one noted several times before, namely, some overlap among the four primaries A, F, H and Q_2(-), marking the second order extraversion factor, but it is only of moderate proportions. One notes that these deviations are generally among the rating variables, which are inherently compound meanings and, in spite of assigned definitions, continue to have somewhat different meanings for different raters. These comments on "fit" seem appropriately made here, where Table 3 is first seen; but actually the final fit is examined precisely by congruence coefficients, below.

As to the *mean* size of loadings of the four variables set out to mark the hypothesized factor they reach, when averaged across the 16 factors: Rating 1, +.41; Rating 2,-.40; Questionnaire A Form, +.50; Questionnaire B Form, +.46. If these were to be treated as equal on the whole to correlations and corrected for attenuation by being divided by the square root of the mean reliabilities in Table 2 they would yield roughly .50 for ratings and .56 for questionnaire scales, leaving little doubt that they are unmistakable markers for the hypothesized factors, and indeed capable of giving pretty good score estimates for those factors when a weighted pool of the four is taken.

It remained to see how far the preliminary identification of these factors as A, Affectia, B, Intelligence, C, Ego Strength, etc., through recognizing the four markers "by eye" is sustained by precise congruence or s coefficient calculation over the whole pattern, i.e., with the lesser loadings on other variables considered. Inasmuch as no previous study — except our recent small 82 sample one (Cattell, Finkbeiner, et al, 1973) which was independent as a factor analysis but not entirely so in sample — has been made with 32L and 32Q variables we have to proceed (1) by comparing the L-data (rating) patterns with previous patterns, (2) comparing the Q patterns with those found when forms A and B (32 variables) of the 16 P.F. have been factored alone, i.e., the matching for our cross media

factors has to be made with the best definition obtained previously in the separate media — an exacting test. The second can be done exactly since several large sample factorings of 16 P.F. scales — 2 to 4 per factor — are available (Cattell, 1973). The first, however, has to be done on slightly fewer variables because in the early researches (Cattell, 1945, 1946, 1947, 1948) the listed variables (Table 17, p. 295, Cattell, 1946) come close enough to identical definition for at most 30 variables. Congruence coefficients on the factor-pattern values are shown in Table 4. The criterion researches used for the matching here are a rating research (Cattell, 1945) meeting the group rating and reliability conditions above and a recent factoring of the 16 P.F. Form A and B scales on a large student sample comparable to the present one (Kulhavy, 1975). By the number of factor tests available at that time, only 12 factors were taken from the ratings, and a modern check indicates 2 or 3 more, but this difference would distort the 12 very little. Although no questionnaire or rating markers were specifically put in for D, J and K, such as might assist the rotation of these factors in the ratings, it is interesting to note that patterns appropriate to D and J (r_c values of .52 and .60 respectively) have emerged clearly in the rating variables (factor 19[-] and 21[-]). (The congruence of K of .41 with Q_2 is suggestive for further inquiry, but one suspects so moderate a value could be a cooperative factor effect.)

Table 4

Checking of Personality Factor Trait Identities

by Congruences with Previous Studies

(a) Life Criteria (L-data) Rating Factors

Present	Identification 1946		r_c with match	r_c of present pattern with highest rival alternative (1946)
1(A)	A	Affectothymia	.47**	.23 with F
2(B)	B	Intelligence	.54**	None significant
3(C)	C	Ego Strength	.50*	.64 with B, .61 with G
19(-)	D	Excitability	.52*	With present Q_4, .56, already better matched.
4(E)	E	Dominance	.68**	-.52 with J
5(F)	F	Surgency	.70**	-.68 with L, .46 with A, .38 with B
6(G)	G	Super Ego	.35*	.46 with A, .38 with B; last two better matched.
7(H)	H	Parmia	.56**	.55 with F, already better matched.
8(I)	I	Premsia	.21*	No better match
21(-)	J	Coasthenia	.60*	-.62 with N, -.72 with Q_2
No match	K	Mature Socialization		.41 with Q_2. Not accepted as match.
9(L)	L	Protension	.25	Better with D, .55
Mean for 10 Accepted Matches:			.52	

Table 4 Continued

(b) Questionnaire (Q-data) Factors

Identification Present	1973*	r_c	Identification Present	1973	r_c
1 (A)	A	.65**	8 (I)	I	.91**
2 (B)	B	.75**	9 (L)	L	.43*
3 (C)	C	.44**	10 (M)	M	.43**
4 (E)	E	.68**	11 (N)	N	.20* Questionable Match
5 (F)	F	.70*	12 (O)	O	.32*
6 (G)	G	.79**	13 (Q_1)	Q_1	.67**
7 (H)	H	.81**	14 (Q_2)	Q_2	.29 Not Matched
			15 (Q_3)	Q_3	.65**
			16 (Q_4)	Q_4	.37 Questionable Match

Mean for 15 Accepted Matches: .59

By the Schneewind (1970) tables all r_c's are significant at $p < .05$ above .2, approximately. Where ** are shown, the figure is also the highest in the r_c matrix row and column. With *, it is the highest in one, but the alternative in other row or column is exceeded for the factor concerned, ruling out the alternative. Decisions to accept these matches are partly based on the salient variable similarity index, s, calculated for borderline instances.

The final inference in a matching procedure cannot depend on the absolute r_c value with the hypothesized match alone. The match can be accepted provided the value is the highest in the column *and* row, i.e., provided it lies on the diagonal of the full r_c matrix, even if it should be of borderline significance ($p < .05$ is about .20 here and $p < .01$ is .30. See Table 4.) in size. At the same time significant values are to be expected with factors long known to be "cooperative," notably among A, H, F and L, and among C, G and Q_3. Such secondary values are shown in Table 4(a) for ratings, but none reached real rivalry in the Q-data (Table 4[b]). Our final matching conclusions considered this total gestalt: the double asterisks show where all conditions were met, i.e., highest in row and column, and $p < .01$ significance.

It is consistent with earlier work (Cattell, 1973) that the weaker matches and possible confusions fall as they do, namely in weaker matches for the smaller variance factors L, N, O, Q_2 (the weaker match for I is exceptional and in ratings only) and in possible confusions of D, L, O and Q_4, of C with G, and mutually of A, F, H and L, the exvia components. But when all considerations are invoked, we conclude that 10 of 12 rating factors and 15 of 16 questionnaire factors are matched. That Q-data results are better would be expected from the fact that L-data "markers" are less salient, by the nature of ratings, from non-markers, and that the cross-identification of the variables themselves has suffered by some re-shaping of their definitions in the thirty year interval.

The Q_2 loss of definition seems due to that absorption of variance into other exvia factors, notably A and H, which expresses the domination of some raters' perceptions by the second order extraversion concept, making separation of the primaries A, F, H and $Q_2(-)$ always less than by test scores. The difficulty evidenced in Table 4(b) — and more clearly in the 16 x 16 table not shown here from which it is condensed — in separating O and Q_4 is also traditional, the separation of these two correlated primaries becoming satisfactory only with the highest levels of technical procedures.

With the exceptions noted, however, the verdict is definitely that the factoring of L- and Q-data *together* leads to the same series of factors, with the matched cross-identifications previously assigned to them in *separate* L- and Q-analyses (Cattell, 1957, 1973). But we can do a more extensive testing of the "indifference of medium" hypothesis than through the matching of factor patterns alone, since the Cattellian theory of personality structure also specifies particular second and third strata structures, which again are postulated to be the same, when freed of instrument factors, seen through the L- and Q-media (Cattell, 1957, 1973). This article cannot give space to those higher order analyses, but the primary factor correlation matrix from which they can be derived has been set out elsewhere (Vogelmann & Cattell, 1976), and the second stratum structure there demonstrated for the present primaries is highly consistent with that of past researchers.

DISCUSSION

The concept of primary personality factors as put forward by Cattell (1946, 1957, 1973) has had to withstand attacks from three main directions. First, criticisms come from those who cannot empirically reproduce the factor patterns themselves; either in one medium or as determiners common to two or more media of observation. Secondly, when the hypothetical pattern structure has been checked by exact methods, there are possible theoretical positions explaining the agreed structure differently. We need not concern ourselves here with the examinations of alternative psychological interpretations, which need to proceed for some time, but only with the almost philosophical issue of whether the discovered patterns reside in language itself, as proposed in different ways by Passini & Norman (1966), by Novakowska (1973) and implicitly by Osgood (1962). These share the position that the factoring of ratings couched in verbal symbols is really nothing more than the dimensions of language. The primary personality structure theory encounters, thirdly, the position of Mischel and some Skinnerians that behavior is so specific in its acquisition that *a priori* one would expect no common trait patterns of any kind. The last speculation is sheer *a priori* and can be put forward only by declining to inspect correlational evidence. Actually, it must be completely dismissed the moment one looks at the substantial communalities in personality correlation matrices, no matter *how* the variables are measured. The second objection is surely sufficiently answered by (a) evidence that the same patterns exist as in verbal ratings and in objective when we go right away from verbal assessment into objective, laboratory tests (Cattell, 1955, 1973), and (b) the demonstration (Dickman, 1960) that factoring observer ratings yields the same factors as are found by factoring objective physical measurements when we go to *physical* systems. In the Cattell-Dickman Ball Problem, for example, the factors of size, weight, elasticity, etc., are positively

known from the outside evidence of physics. The general conclusion must surely be that in sane subjects, as Korzibsky argued, the "cognitive map" as expressed in verbal usage, corresponds to the structure in the external world. It comes to do so, by trial and error learning (or maybe as Chomsky argues, partly by genetic adaptation). This surely is a more parsimonious explanation of the Passini & Norman results than some others offered.

There remains the first type of objection, and here one must reluctantly point to grave errors and omissions in published factorings failing to find the common patterns. At least three experienced researchers in this field have taken the trouble to pick up these maverick studies and show where several necessary steps in factor analytic experiment have invariably been omitted by those who finish with unrecognizable factors, radically different in number or nature from those replicated patterns on which systematic, programmatic research persistently converges (Cattell, 1973; Digman, 1963; Vaughan, 1973).

We reach at this point, therefore, the position that factors A through O, and possibly through Q_6, can be replicated in the separate media, in terms of general psychological meaning, and the issue remaining is whether they match, as expected, by objective correlation of patterns, across the media. One must recognize that there exist some prominently quoted but little technically examined studies which do reach the usually replicated primaries, but then show quite disappointingly low correlations between scores on the same factors in two different media. The finding of low correlations between the factor scores of individuals for supposedly the same factor in two media is, however, quite different from a failure of alignment of factor patterns when a common factoring is made simultaneously of data from two media, as here. Reports of low correlations between variables from which a given factor is scored in two different media occur in the early article of Becker (1960), in Bohleber (1964), and the careful work of Nesselroade & Schaie (1973). Actually, however, as the two last point out, there is no problem in the low correlations as such; it is only created in the interpretations made of them. If an L-data and a Q-data marker each mark the factor 0.6 — which is enough, with other markers to leave no doubt about a factor pattern — still they will not show a correlation of factors so estimated exceeding the rather miserable value of 0.36 (i.e., 0.6 x 0.6) as pointed out above. This is no disproof of the common factor in L- and Q-data, though admittedly for practical test purposes we should work hard to find further variables that will bring correlations of the same factor estimated in two media to more impressive levels. In seeking further clarification of trait meaning, by discovering markers with higher loading values, the great need is to remedy the widespread weakness evident in past rating procedures themselves. For example, in both cases of low relations just cited, the children were rated by teachers, i.e., in a narrow situation, and most studies have yet other defects of rating which have grossly reduced the validity and reliability of variables as such. Even in that situation, however, with the statistical estimates which Nesselroade made in correlation of Schaie's data, the factor analytic results pointed in the main to the present conclusion of a one-to-one relation of personality source traits as perceived through the two media.

Apart from the main conclusion on matching above, three other findings of this study are of interest to ongoing programmatic research in this area: (1) further support is given to the general conclusion (Cattell, 1973) that there are as

many as 23 or possibly more primaries in any good sampling of the personality sphere. This support for a personality dimensionality of not less than 23 defines a position obviously remote in the hypothesized number and the consistency of replication from the mutually contradictory theories of 2, 3, 5, 6, 11 and 15 in the writings respectively of Peterson (1965), Eysenck & Eysenck (1963), Norman (1961), Tupes & Cristal (1961), Sells, Demaree & Will (1971), and Howarth & Browne (1972), though agreeing closely with the recent work of Dermen (1973) with Harman.

(2) The four factors Q_1, Q_2, Q_3 and Q_4 which had been kept with a Q number notation rather than an alphabetical order, because for twenty years since confirmation in questionnaires (1949, 1950) they had remained unmatched in L-data, are now shown to have L-variance on rating marker variables. In fact, as Table 3 shows, very cleancut pairs of rating marker variables can be found for each, and they agree with the psychological meaning long given to Q_1, Q_2, Q_3 and Q_4 (Cattell, Eber and Tatsuoka, 1970), from the nature of their "mental interiors" and their criterion associations. However, for the present they will be left with Q designation indexing only.

(3) Some light is thrown on the question of instrument factors. Although this was not our main objective it is obvious that the analysis of common L & Q data on the same subjects presents a more ideal avenue to detecting instrument factors than by attempting to make all inferences from analysis of Q-data alone, as in virtually all instances surveyed by Wiggins (1973). By hypothesis, and present knowledge of such factors (Cattell, 1973) it is extremely unlikely that an instrumental factor will stretch across more than one medium. Although trait view theory (Cattell, 1968) tells us that instrument factor distortions in L-data could be as powerful as in Q, it points to their appearing only in *a single rate* and with reference to his role and personality factors. In the present study 8 or 9 people rated each person, which should be enough to cancel out all L-data instrument intrusion except to the most sensitive statistical treatment. On the other hand, the possible "desirability" instrument factors in the questionnaire should remain and they should stand out by presenting us with factor loading patterns which (a) appear only in Q-data and (b) have some correspondence to the Q-data factors which are known to change significantly in level through desirability distortion, e.g. the contrast of anonymous and signed 16 P.F. results (Cattell, 1973).

Two factors with such characters appear here — No. 19 & 20 in Table 3. The loadings are not high but they are consistent in contrasting Q loadings that are significant with L loadings insignificant or mildly opposite in sign. Thus 19 has .36 and .32 on Q markers on A and .04 and -.13 on L markers (one must remember that the second L marker in Table 3 always has to be reversed in sign). On F it has .77 and .60 against .01 and -.13 while on E, H and (less definitely) Q_1 it has indications of the same L and Q discrepancy. On No. 20 the powerful difference effects are on G (.65 and .22 versus .01 and -.20), and on Q_3 (.69 and .15 versus .01 and -.14). The total set of these two series of primaries A, C, F, H, M(-), Q_3 and Q_4 (E is anomalous), constitute the main factors on which shift occurs, consistently in the directions here, when subjects "fake good" (Cattell, 1973, Table 56, p. 393). It is conceivable that a larger sample and equally careful rotation would reveal some other exclusively Q-data factors with loadings concordantly on socially desirable primaries. But all factors beyond the

well known 16 P.F. ones here except the above are either trivial, as No. 22 essentially is, or have *combined* loadings on L and Q as do Nos. 21 and 23 and presumably represent new real factors.

While we would thus not definitely subscribe to a two factor theory of desirability distortion, it appears that two factors account for most, and by their natures they can distinctively be seen as "a normal good fellow" target, mainly on extravert and adjustive factors — A, F, Q_4(-) etc. — and a "morally superior" pattern in 20, loading super ego, G and self-sentiment, Q_3. The latter some might call the "hypocrite factor" dealing with both moral pretentiousness and pretentiousness in social repute, as in a "stuffed shirt" superiority.

Granted that these two are the major factors with a suitable character for desirability distortion, then the overall results also strongly support the present writer's conclusion that the dimension of desirability vs. self depreciation present in the anxiety factor, Q_{II} (No. 3 in Table 3) is a real, instrument-free trait factor, but with items which necessarily have social undesirability secondarily attached to them. The present writer has protested (1973, and with Eber and Tatsuoka, 1970) that a proposed practice of throwing out all desirability-undesirability variance as if one big "desirability" factor (It does not conform to one big factor anyway!) is a bad case of throwing away the baby-with the bath water. A substantial part of undesirability variance is created by the anxiety factor, and we have clear evidence in the second order factoring of the present data (Cattell & Vogelmann, 1976) that Q_{II}, Anxiety, is as solid in observer rating as it is in Q-data.

The contribution of the present research to the faking, and "desirabiity" investigations, therefore, is to the effect that much of the undesirability present in the variance of the anxiety factor is an unescapable part of this true trait variance and not an instrument factor artefact; but that two instrument factors exist — contingently designatable as the "good fellow" and the "moral impeccability" instrument factors — which are (conscious or unconscious) directions of desirability faking. This interpretation is supported by the next higher order analyses of the present 23 factors (Cattell and Vogelmann, 1976) where we find the desirability loaded, negatively, on this "good upbringing" factor, Q VIII. According to trait view theory this extracted factor, as contrasted with the performance on actual scales, should be the *real* personality factor, instrument-face; and what this loading is telling us is that the genuinely more moral person (good upbringing in high G, super ego and high Q_3, self-sentiment) employs both of these faking tendencies *less* than the less moral.

While the above coheres, we must not overlook the fact (Table 4) that No. 19 has a significant correlation (in the "non-good-fellow" direction) with the pattern of correlation with ratings which seem to characterize the D factor of "excitability" (Higher C, E, and Q_4 markers in ratings). That is to say "good fellow" distortion *might* yet prove to be part of a real personality factors, D(-).

(4) Factors 17, 18 and 22 here are factors also almost wholly in Q-data. They do not, however, make sense, as instrument factors directly and should be investigated in future as possibly being the "missing factors" scrutinized and checked in the Cattell and Delhess (1973), Cattell and DeVoogd (1972) and Watterson and Cattell (in press) researches. Since no markers for these were yet available for inclusion this hypothesis obviously cannot be checked here, and

the same applies to Factor 23 which, although convincing in its loadings in both media, could be one of the missing factors — Q_5, Q_6 and Q_7 — yet unaccounted for. Actually, No. 17 is so feeble in loadings as to be best put at the end of this series as probably an error factor, and, if so, this would leave three significant factors here unidentified — No. 18, 22 and 23 — and three known in previous research — Q_5, Q_6 and Q_7 — but lacking inclusion of their markers here. Alternately one could ask if No. 23 is an instrument factor in observer ratings, but the consideration of rating procedures above make this unlikely.

At various levels of probability set out above, therefore, we find some 17 factors, spanning both media, and matching trait patterns that have been already assigned names and symbols in general research. After these 17 there are 2 hypothesized to be instrument factors in Q-data (though one of these, No. 19, *could* be D), then 3 unidentified awaiting possible matching with Q_5, Q_6 and Q_7, and finally one formless error factor. The writers wish to thank the students in Oregon and Illinois undergraduate classes for their conscientious cooperation in a long experiment.

REFERENCES

1. Becker, W. C. The matching of behavior rating and questionnaire personality factors. *Psychological Bulletin,* 1960, *57,* 201-212.
2. Bohleber, M. E. Correlates of social sensitivity among elementary school teachers. Wisconsin Teacher Education Project. NIMH Grant 2M-6624. Report, 1964.
3. Buss, A. R. A general developmental model for inter-individual differences, intra-individual differences, and intra-individual changes. *Developmental Psychology,* 1974, *10,* 70-78.
4. Campbell, D. T., and Fiske, D. W. Convergent and discriminant validation by the multi-trait multi-method matrix. *Psychological Bulletin,* 1959, *56,* 81-105.
5. Cattell, R. B. The description of personality. Principles and findings in a factor analysis. *American Journal of Psychology,* 1945, *58,* 69-90.
6. Cattell, R. B. Confirmation and clarification of primary personality factors. *Psychometrika,* 1947, *12,* 197-220.
7. Cattell, R. B. The primary personality factors in women, compared with those in men. *British Journal of Psychology,* 1948, *1,* 114-130.
8. Cattell, R. B. Psychiatric screening of flying personnel. Personality structure in objective tests. USAF School of Aviation Medical Report No. 9, 1955 (Project No. 21-0202-0007), 1-50.
9. Cattell, R. B. *Personality and motivation structure and measurement.* New York: World Book, 1957.
10. Cattell, R. B. The configurative method for surer identification of personality dimensions notably in child study. *Psychological Reports,* 1965, *16,* 269-270.
11. Cattell, R. B. Trait view theory of perturbations in ratings and self-ratings (L, BR, and Q-data): Its application to obtaining pure trait score estimates in questionnaires. *Psychological Review,* 1968, *75,* 96-113.
12. Cattell, R. B. Comparing factor trait and state scores across ages and cultures. *Journal of Gerontology,* 1969, *24,* 348-360.
13. Cattell, R. B. *The description and measurement of personality.* New York: World Book, 1946. Reissued Johnson Reprint Co., 1969.

14. Cattell, R. B. *Personality and mood by questionnaire.* San Francisco: Jossey-Bass, 1973.
15. Cattell, R. B., Balcar, K. B., Horn, J. L., and Nesselroade, J. R. Factor matching procedures: An improvement of the *s* index, with tables. *Educational and Psychological Measurement,* 1969, *29,* 781-792.
16. Cattell, R. B. and Delhees, K. Seven missing normal personality factors in the questionnaire primaries. *Multivariate Behavioral Research,* 1973, *6,* 173-194.
17. Cattell, R. B. and DeVoogd, T. A check on the structure of the comprehensive 23 personality primaries in Q-data. Advanced Publication No. 28, I.R.M.A.S., Boulder, Colorado, 1974.
18. Cattell, R. B. and Digman, J. M. A theory of the structure of perturbations in observer ratings and questionnaire data in personality research. *Behavioral Science,* 1964, *9,* 341-358.
19. Cattell, R. B., Eber, H. W. and Tatsuoka, M. *Handbook for the 16 P.F. Questionnaire.* Champaign, Illinois: Institute for Personality and Ability Testing, 1970.
20. Cattell, R. B., Finkbeiner, C., Willes, P., Robertson, J., and Brim, B. Common factoring of observer ratings and questionnaire measures on a sample of 75 young adults. Advanced Publication No. 30, I.R.M.A.S., Boulder, Colorado, 1973.
21. Cattell, R. B., Horn, J. L., Sweney, A. B. and Radcliffe, J. A. *The motivational analysis test.* Champaign, Illinois: Institute for Personality and Ability Testing, 1964.
22. Cattell, R. B. and Jaspars, J. A. A general plasmode (No. 30-10-5-2) for factor analytic exercises and research. *Mult. Behav. Res. Monographs,* 1967, 67-3, 1-212.
23. Cattell, R. B. and Saunders, D. R. Beitrage zur Faktoranalyse der Personlickkeit. *Ztschrift f. exper. u. angewandt. Psychol.,* 1955, 325-357.
24. Cattell, R. B. and Vogelmann, S. Second stratum personality structure in joint rating and questionnaire measures; with new light on questionnaire distortion. *Multivariate Experimental Clinical Research,* 1976, 2, 1, 43-56.
25. Cattell, R. B. and Vogelmann, S. A comprehensive set of trials of the scree test criterion for determining the number of factors. *Mult. Behav. Res.,* 1977, 10, 1-17.
26. Coan, R. W. and Cattell, R. B. Child personality structure as revealed in teacher behavior ratings. *Journal of Clinical Psychology,* 1957, *13,* 315-327.
27. Dermen, D. A verification study of 28 self report personality factors. In H. H. Harman (Ed.) *Technical Report No. 3,* 1973, Princeton, New Jersey: Educational Test Service.
28. Dickman, K. W. Factorial validity of a rating instrument. Ph.D. dissertation. University of Illinois Library, Urbana, Illinois, 1960.
29. Digman, J. The principal dimensions of child personality as inferred from teachers' judgments. *Child Development,* 1963, *34,* 43-60.
30. Eysenck, H. J. and Eysenck, S. B. G. *Manual for the Eysenck Personality Inventory,* San Diego: Educational and Industrial Testing Service, 1963.
31. Fiske, D. W. *Measuring the concepts of personality.* Chicago: Aldine, 1971.
32. Howarth, E. and Browne, J. A. Investigation of personality factors in a Canadian context. *Canadian Journal of Behavioral Science,* 1972, *4,* 85-90.

33. Hundleby, J., Pawlik, K., and Cattell, R. B. *Personality factors in objective test devices.* San Diego: Knapp, 1965.
34. Kaiser, J. F., Hunka, S., and Bianchini, J. C. Relating factors between studies based upon different individuals. *Multivariate Behavioral Research,* 1971, *6,* 409-422.
35. Kristof, W. Testing a linear relation between true scores of two measures. *Psychometrika,* 1973, *38,* 101-111.
36. Kulhavy, R. and Cattell, A. K. S. *The structure and validity of the 1972 edition of the 16 P.F.,* in press.
37. Lord, F. M. and Novick, M. *Statistical theories of mental test scores.* Reading, Mass.: Addison-Wesley, 1968.
38. Marshall, D. and Cattell, R. B. *The seven scale supplement to the 16 P.F.* Champaign, Illinois: Institute for Personality and Ability Testing, in press.
39. Nesselroade, J. R. and Reese, H. W. *Life span developmental psychology.* New York: Academic Press, 1973.
40. Norman, W. T. Development of self report tests to measure personality factors identified from peer nominations. USAF Technical Report No. 61-44, 1961.
41. Novakowska, M. The limitations of the factor-analytic approach to psychology, with special application to Cattell's research strategy. *Theory and decision,* 1973, *4,* 109-139.
42. Osgood, C. E. Studies on the generality of affective meaning systems. *Amer. psychol.,* 1962, *17,* 10-28.
43. Passini, F. T. and Norman, W. T. A universal conception of personality structure. *Journal of Personality and Social Psychology,* 1966, *4,* 44-49.
44. Peterson, D. R. Scope and generality of verbally defined personality factors. *Psychological Review,* 1965, *72,* 48-59.
45. Peterson, D. R. and Cattell, R. B. Personality factors in Nursery school children as derived from parents' ratings. *Journal of Clinical Psychology,* 1958, *14,* 346-355.
46. Schaie, K. W. On the equivalence of questionnaire and rating data. *Psychological Reports,* 1962, 10, 521-522.
47. Schneewind, K. A. and Cattell, R. B. Zum Problem der Faktoridentifikation. *Psychologische Beitrage,* 1970, *12,* 214-226.
48. Sells, S. B., Demaree, R. B. and Will, D. P. Dimensions of personality. 1. Conjoint factor structure of Guilford and Cattell trait markers. *Multi. Behav. Res.,* 1970, *4,* 391-422.
49. Tupes, E. C. and Cristal, R. C. Recurrent personality factors based on trait ratings. USAF, ASD, Technical Report No. 61-97. Lackland AFB, Texas: Personnel and Training Center, 1961.
50. Vaughan, D. S. The relative methodological soundness of several major personality factor analyses. *Journal of Behavioral Science,* 1973, 1, 305-313.
51. Watterson, D. and Cattell, R. B. Progressive rectification of the nature of the seven new normal factors in the 16 P.F. supplement, in press.
52. Wiggins, J. S. Convengences among stylistic response measures from objective personality tests. Educ. & Psychol. Meas., 1964, 24, 551-562.
53. Wiggins, J. S. Personality and prediction; principles of personality assessment, London: Addison-Wesley, 1973.

NOTES

[1]The marker variable matches used here are (with the present research number, from Tables 1 [first column] and 3, placed first in each pair, and the second identical with that in Cattell, 1945, Table 3, p. 88):

1 = 9(-); 2 = 14; 5 = 2; 6 = 12(-); 9 = 4(-); 10 = 20; 13 = 1; 14 = 7(-); 17 = 15; 18 = 33(-); 21 = 11; 22 = 28; 25 = 26(-); 30 = 27(-); 33 = 19; 34 = 18(-); 37 = 34; 38 = 25(-); 41 = 35; 42 = 3(-); 45 = 29(-); 46 = 32; 53 = 31; 54 = 10(-); 58 = 17; 61 = 18; 62 = 26(-); making 28 variables in common.

Multivariate Experimental Clinical Research, Volume 2, Number 2, 1976, 89-92

A CHECK ON THE SECOND-ORDER FACTOR STRUCTURE OF THE HIGH SCHOOL PERSONALITY QUESTIONNAIRE

Thomas W. Klein
University of California at Davis

ABSTRACT

A second order factor analysis of the HSPQ, using the promax method followed by blind hand rotation to simple structure, was conducted. The results obtained for both males and females replicate well six factors previously reported.

INTRODUCTION AND METHOD

Friel (1972) undertook a large scale check of the primary factor structure of the High School Personality Questionnaire (HSPQ). He administered the HSPQ to 1836 junior and senior high school students from the Harrison, Wetzel, and Wood County, West Virginia school systems. Male and female samples were then divided, at random, into two groups each (456 and 466 males; 410 and 504 females). Data from each of these samples were factored and the solutions compared. This analysis indicated high congruence across samples and good agreement with the fourteen factors reported by Cattell and Cattell (1969). The correlations among the primary factors in each of the four groups agreed well with those reported by Cattell and Cattell, though their values were generally somewhat higher (Friel, 1972). The object of this analysis was to check the second-order factor structure of the HSPQ in this large body of data.

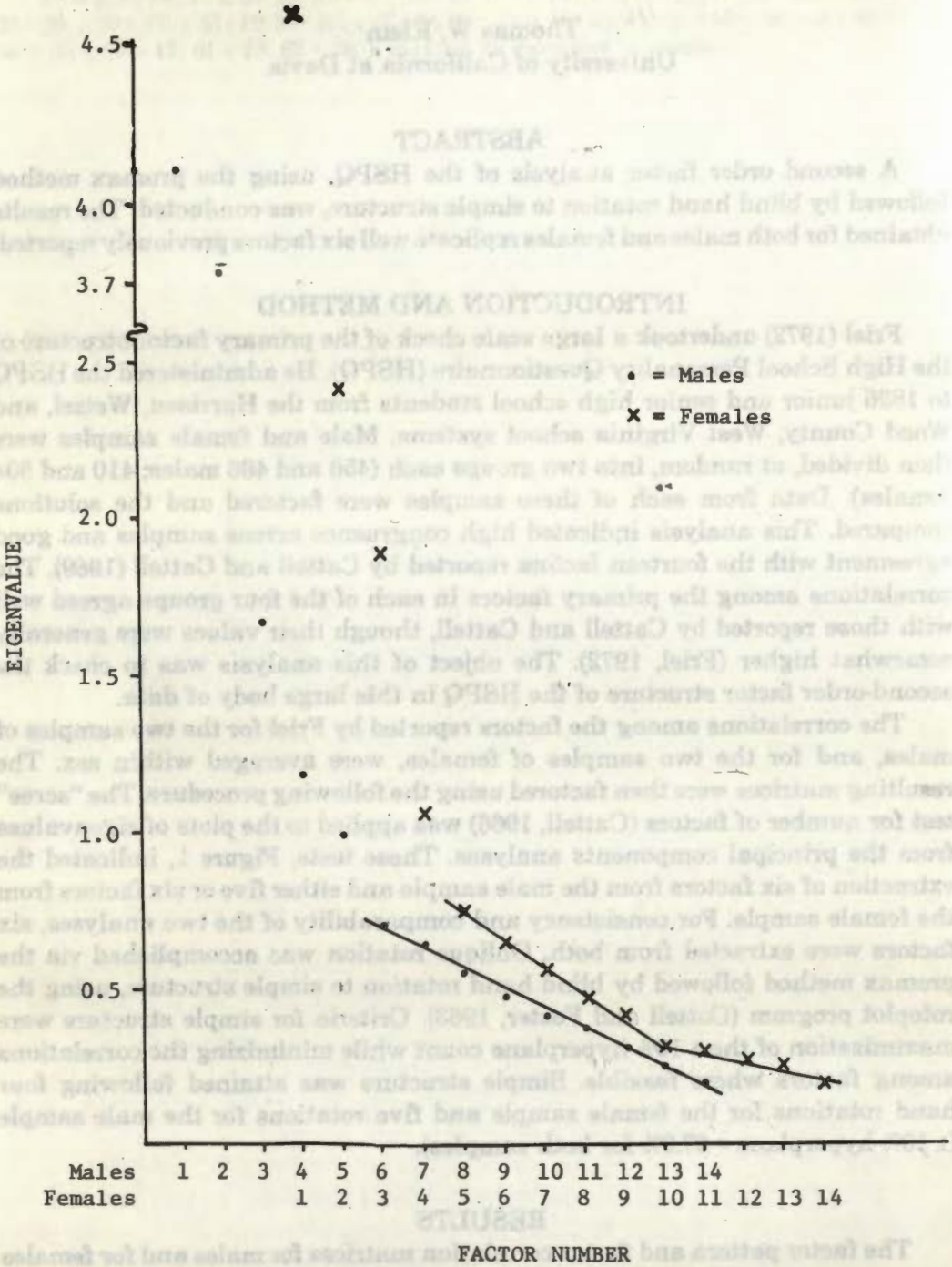
The correlations among the factors reported by Friel for the two samples of males, and for the two samples of females, were averaged within sex. The resulting matrices were then factored using the following procedure. The "scree" test for number of factors (Cattell, 1966) was applied to the plots of eigenvalues from the principal components analyses. These tests, Figure 1, indicated the extraction of six factors from the male sample and either five or six factors from the female sample. For consistency and comparability of the two analyses, six factors were extracted from both. Oblique rotation was accomplished via the promax method followed by blind hand rotation to simple structure, using the rotoplot program (Cattell and Foster, 1963). Criteria for simple structure were maximization of the ± 10% hyperplane count while minimizing the correlations among factors where feasible. Simple structure was attained following four hand rotations for the female sample and five rotations for the male sample (± 10% hyperplane = 67.9% for both samples).

RESULTS

The factor pattern and factor correlation matrices for males and for females are presented in Table 1. Cattell (1973) rotated seven previous second order solutions, including the two reported here, to similar factor patterns. The seven resulting factor patterns were then averaged. These patterns replicated factors I through IV, and factor VI, as reported by Cattell and Cattell (1969), and also yielded a new factor VIII, which consists of -F, G, and Q_3. These roman numerals

Figure 1

SCREE TEST FOR NUMBER OF SECOND ORDER FACTORS IN THE HSPQ, FOR MALE (•) AND FEMALE (x) SAMPLES.



are used for factor identification in Table 1 with the designation VII (Factor B in Cattell, *et al.*, 1970) substituted for VI (factor B in Cattell and Cattell, 1969). Since the order of factor extraction differed in the two samples, the patterns are presented in order by Roman numeral, for ease of comparison.

Factor matching was accomplished through the use of congruence coefficients between the obtained patterns (M and F) with the pattern obtained by Cattell (1973), and with the patterns reported by Cattell and Cattell (1969) for males and for females. These congruences are presented in Table 2.

Table 1

Second Order Factor Patterns for Males and Females on the HSPQ, Their Order of Extraction, and Factor Correlations

Personality Factors	Males I	Males II	Males III	Males IV	Males VII	Males VIII	Females I	Females II	Females III	Females IV	Females VII	Females VIII
A	.67	-.00	-.25	.08	.10	.02	.76	-.12	-.20	-.17	-.04	-.04
B	-.11	-.02	.05	.05	.70	.04	-.03	.03	-.01	-.01	.72	.11
C	.03	-.76	.10	-.09	-.22	-.00	.12	-.53	.09	.10	-.10	.39
D	.04	.85	.10	.06	.00	.04	.11	.91	.10	.10	-.04	-.04
E	-.10	-.10	.56	.45	-.09	.06	.02	-.08	.66	.10	.05	-.06
F	.46	-.05	.05	.34	.16	-.54	.58	.10	.31	.08	.10	-.46
G	.05	-.06	-.09	-.24	.09	.60	-.04	-.09	-.30	-.03	.11	.69
H	.58	-.32	-.01	.10	.08	.02	.67	-.36	.10	.11	-.08	.11
I	-.10	.01	-.63	-.05	-.29	.10	.10	-.05	-.70	.10	.06	.07
J	-.59	.06	.04	.30	.06	-.00	-.08	.09	-.02	.78	-.00	-.10
O	-.22	.53	-.27	.08	-.10	-.11	-.26	.41	-.09	.06	.08	-.33
Q_2	-1.01	-.50	-.09	.09	.10	.00	-.61	-.11	.10	.36	.07	.09
Q_3	-.00	-.08	.04	.04	-.04	.82	-.02	-.47	-.25	.07	-.08	.49
Q_4	.03	.84	.06	.00	-.01	.04	-.05	.79	-.04	-.00	-.10	.01
						Factor Correlations						
I	1.00	-.46	-.03	.10	.12	.13	1.00	-.25	.19	-.33	.23	.25
II		1.00	.00	-.16	-.12	-.52		1.00	-.14	.09	.08	-.47
III			1.00	-.11	-.12	-.17			1.00	.15	-.15	-.08
IV				1.00	-.03	-.11				1.00	-.01	-.03
V					1.00	.16					1.00	.07
VI						1.00						1.00

With the exception of factor IV for females, the congruences with the pattern as obtained by Cattell (1973) are quite good. The factors reported by Cattell and Cattell (1969) also show good replication.

TABLE 2

Congruence Coefficients Between the Solutions Presented in Table 1 (M vs. F) and Those Presented by Cattell (1973), and Cattell and Cattell (1969).

	M vs F	M-C	F-C	M-CC	F-CC
I	.88	.94	.91	.81	.92
II	.92	.89	.80	.81	.95
III	.87	.96	.89	.79	.72
IV	.53	.81	.43	.47	.82
VII	.81	.90	.94	.82*	.64*
VIII	.86	.97	.91	—	—

*Factor VI in Cattell and Cattell (1969)

REFERENCES

1. Cattell, R. B. The scree test for the number of factors. *Multivariate Behavioral Research,* 1966, *1,* 140-161.
2. Cattell, R. B. *Personality and mood by questionnaire.* San Francisco: Jossey-Bass, 1973, p120.
3. Cattell, R. B., and Cattell, M. D. L. *Handbook for the High School Personality Questionnaire "HSPQ".* Champaign, Ill.: Institute for Personality and Ability Testing, 1969.
4. Cattell, R. B., Eber, H. W., and Tatsuoka, M. M. *Handbook for the Sixteen Personality Factor Questionnaire (16 PF).* Champaign, Ill.: Institute for Personality and Ability Testing, 1970.
5. Cattell, R. B., and Foster, M. J. The rotoplot program for multiple single-plane visually-guided rotation. *Behavioral Science,* 1963, *8,* 156-165.
6. Friel, J. C. Replication and cross validation of primary source traits measured by the High School Personality Questionnaire. Unpublished Master's Thesis, West Virginia University, 1972.
7. Friel, J. C., and Nesselroade, J. R. Test of the replicability of Cattell's HSPQ factor structure using item parcels. Paper presented at the meeting of the Eastern Psychological Association, Washington, D. C., May, 1973.

NOTE

[1]I thank Mr. John Friel for his kind permission to use his data in this analysis.

Multivariate Experimental Clinical Research, Volume 2, Number 2, 1976, 93-100

FACTORS OF SUBORDINATE-SUPERORDINATE INTERACTION[1]

Arthur B. Sweney
Leslie A. Fiechtner
Wichita State University

ABSTRACT

Two hundred-four crew members and sixty-nine crew commanders of a Titan II Missile Wing completed a battery of instruments measuring superordinate subordinate role preferences, pressure, and perceptions. After matching subordinate variables with the appropriate superordinate variables, the data were correlated, factored, and rotated for orthogonal simple structure. Of the twelve factors obtained five were identified inclusively with crew commander behavior; three were related to the behavior of the crew members; and four described the interactions between the two realms. Perceptions were found to be primarily a function of observer variance and the roles were found to have much less shaping effect across levels than was expected.

INTRODUCTION

The literature in group dynamics and in the behavioral science applied to organizational behavior is rich in its implications concerning the interactions between behaviors and various hierarchical levels, but it is very limited in reports of explicit studies in this area.

The Ohio State studies of management behavior stated by Hemphill (1950) and continued by Stodghill et al (1962) studied perceptions and expectations of subordinates concerning their superordinates, but did not make any effort to gather data directly from both levels. McGregor did not try to empirically validate his "Theory X" and "Theory Y" assumptions on any level. If he had, he might have had to face up to the reconciliation of conflicting perceptions which it would have entailed.

The closest approximation available to an interactive system is the contingency model by Fiedler (1962). Even this one does not define expected relationships between subordinates and superordinates except in the gross terms of "favorable" and "unfavorable" conditions. In most cases, these terms are reserved for the task characteristics and the physical setting rather than the social interaction.

To fill this need, Sweney (1970, 1971) presented his *Response to Power Model* to establish a role interaction basis for studying the relationships established between superordinates and subordinates. To facilitate research, he developed five instruments to measure various aspects of the six roles: authoritarian, equalitarian, permissive, rebel, critic-cooperator, and ingratiator. The *RPM* model is illustrated in Figure 1.

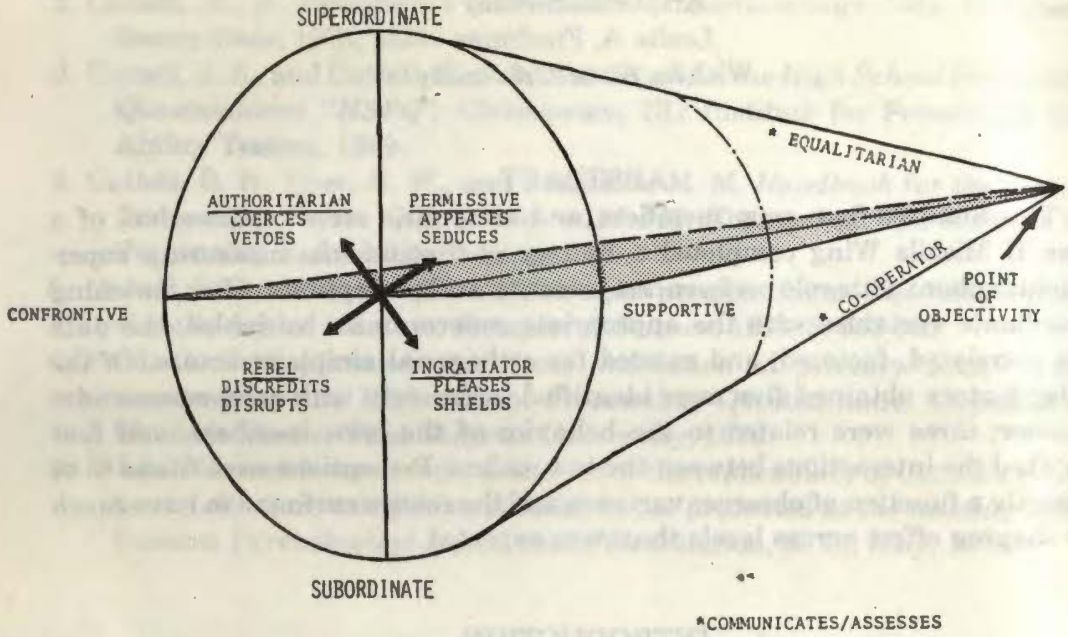
PROCEDURES

The study was conducted in the 581st Missile Wing with the instruments to measure the *RPM* model.

RESPONSE TO POWER MODEL

Figure 1

SIMPLE PERSPECTIVE



INSTRUMENTS

The battery of tests included three self report instruments and two rating scales. The *Response to Power Measure* (RPM) is a scale on preference instruments utilizing ninety-six opinionnaire type items to measure the basic values placed upon the six subordinate-superordinate roles.

The *Supervise Ability Scale* (SAS) and the *Responsibility Index* (RI) were developed to measure social desirability pressure to play certain roles as superordinates or subordinates. Each of these two instruments is self ipsatized and hence has negatively correlated scales. Thus, they measure the distribution of superordinate or subordinate role pressure but not the total amount.

The *Supervisor's Role Rating* (SRR) and *Subordinate Behavior Rating* (SBR) are rating scales of superordinates and subordinates by the member of the opposite level. They have ten and twelve three-response items respectively and are scored to reflect the three superordinate roles and the subordinate roles.

SUBJECTS

The subjects for the studies were drawn from the two operations squadrons of the 381st Missile Wing. These units are each composed of thirty-six four-man crews who served extended tours of duty together in missile silos on a schedule involving 24 hours on duty and 48 hours off duty. Thus, they shared an intimate experience with each other which should lend itself to interactional change and to firm perceptions of each other.

ADMINISTRATION

The subjects from each squadron were tested during one of four pre-departure briefings. They met as a group with adequate spacing to insure privacy of

response. The complete battery took an hour. The crew commanders and his subordinates took the same basic battery except for the *Superviseability Scale* and the *Subordinate Behavior Rating* which were not appropriate at the lower levels.

ANALYSIS

After scoring the test, results were arranged in a data matrix in which each subordinate served as an observation; next to his own test data appeared the test data for his particular crew commander. In this way it was possible to obtain correlations between test scores of crew members and crew commanders as well as the relationships within each of these groups. Thus, the data array consisted of thirty-four variables gathered on 204 subordinates and 69 superordinates.

The correlation matrix was factored using a principle component extraction and a Varimax Rotation. The latter was selected over an oblique rotation because of the exploratory nature of the study and the desire for independent factor variance assessments of some of the variables being studied. The selection of the number of factors was not a crucial decision so the fairly quick but conservative "Guttman criterion" was applied.

RESULTS

Twelve factors were extracted before the eigenvalue dropped below 1.00. This also represented a discontinuity in decrements which added credibility to the assumption that this was also a reasonable place to discontinue factoring.

The rotated factors obtained are disclosed in Table 1. Only loadings out of the .20 hyperplane have been identified. On this basis, five of the factors were predominantly related to superordinate variables, three with subordinate variables and four with interactions between the two realms.

Factor 1. This describes an authoritan vs equalitarian role preference on the part of the crew commanders. It has no ramifications on the subordinate level so it can be assumed to be primarily internalized and irrelevant to the subordinate-superordinate interaction.

Factor 2. This describes an authoritarian vs equalitarian role pressure experienced by the crew commanders. It is augmented by small authoritarian ingratiator loadings from the preference areas. Only one small loading for subordinate behavior was found so that it must be assumed that this also has only minimal or mixed effects upon the interaction process.

Factor 3. This factor involves superordinate true saying response style on the only instrument on which it could be registered. High true saying and low indecisiveness is associated with authoritarianism preference and lack of permissive pressure. No subordinate variables had appreciable loadings and hence this response style was perceived as having limited impact upon the social situation, and could not be associated with any effective ongoing reenforcement treatment by the superordinate.

Factor 4. This factor organizes the subordinate role pressure behaviors of the per crew commanders with their perceptions of their own superordinates (site commanders.) Rebels see their commanders as authoritarians and critics see their commanders as equalitarian. Some slight effect is felt on the level of their subordinates who tend to imitate their roles.

Factor 5. This factor involves the role pressure of crew commanders for rebelliousness vs ingratiation. It shows that the permissiveness both as a role pressure and role preference are associated with ingratiation. Again only negligible loadings are found on the part of the subordinates. These again suggest trends toward imitative behavior.

Factor 6. This factor organizes the superordinate role preferences of the crew members with their response styles. As on the upper level, true saying is positively related to authoritarianism and negatively related to equalitarianism. No meaningful loadings were found on the superordinate level; so this response style behavior can be considered as having little effect upon the social interaction as measured by the test results.

Factor 7. This is an upward perception of permissiveness by crew members of their crew commanders. Low questioning and low critic and ingratiation role preference seem to be related to the process. Some small loadings on superordinate permissiveness indicates that the perceptions are partially justified but primarily the products of the perceivers' own role needs.

Factor 8. This factor relates subordinate role preferences and role pressures of the crew members with their perceptions of their crew commanders. Rebels see their commanders as authoritarians, and ingratiators see their commanders as equalitarians. There was no direct support for their perception from any of the superordinate role measures of the crew commanders. Again, there were small indications that the subordinate roles were a function of imitation for the crew commanders own subordinate role.

Factor 9. This factor shows a clear interaction between superordinate and subordinate levels. The crew commanders of the critic (favorable) subordinate roles for the crew members seem to be a reflection of the subordinates avoidance for the permissive or ingratiating roles. This suggests that in the close confines of the missile silo, artifice was rather quickly discovered. It is interesting that the role perceptions did not bear any direct resemblance to self described roles of the subordinates.

Factor 10. This factor reflects the tendency of the crew commanders to view their subordinates as either rebel or ingratiating. This is strongly related to actual rebel role preference on the part of the subordinate. These same crew members seemingly failed to prefer an equalitarian role style for themselves.

Factor 11. This factor integrates the upward perceptions of the two levels. Crew commanders who see their bosses (site commanders) as permissive are seen as authoritarians by their subordinates. Those who perceive their site commanders as authoritarians are seen as equalitarians by their crew members. This indicates that the best self measure of a superordinate role is obtained from the rating the person makes of others.

Factor 12. This factor indicates that supportive behavior on the upper level is matched by supportive behavior from the lower level. This allows the subordinates to be less equivocal as measured by the use of the middle response on the *RPM* and to exercise less authoritarian role preference.

TABLE 1
SALIENT LOADINGS OF THE FACTOR MATRIX

SUPERORDINATE FACTORS

Instrument	Scale	Loading
	Factor I	
RPM	Authoritarian (Super)	+.62
RPM	Equalitarian (Super)	-.79
RPM	Critic (Super)	-.22
RPM	Rebel (Super)	+.79
RPM	# of Trues (Super)	+.33
SAS	Permissive (Super)	-.21
	Factor II	
RPM	Authoritarian (Super)	+.35
RPM	Ingratiator (Super)	+.24
SAS	Authoritarian (Super)	+.93
SAS	Equalitarian (Super)	-.78
SAS	Permissive (Super)	-.33
RPM	Rebel (Sub)	+.21
	Factor III	
RPM	Authoritarian (Super)	+.43
RPM	# of Trues (Super)	+.83
RPM	# of Questions (Super)	-.91
SAS	Permissive (Super)	-.21
	Factor IV	
RPM	Critic (Super)	-.34
RPM	Authoritarian (Super)	+.81
SRR	Equalitarian (Super)	-.87
SAS	Equalitarian (Super)	-.34
SAS	Permissive (Super)	+.36
RI	Rebel (Super)	+.40
RI	Critic (Super)	-.63
RPM	Rebel (Sub)	-.26
RI	Critic (Sub)	-.30
	Factor V	
RPM	Permissive (Super)	+.27
SAS	Equalitarian (Super)	-.28
SAS	Permissive (Super)	+.52
RI	Rebel (Super)	-.80
RI	Critic (Super)	+.29
RI	Ingratiator (Super)	+.86
RPM	Rebel (Sub)	-.24
RI	Ingratiator (Sub)	-.26

NOTE: (Super) = Superordinate
(Sub) = Subordinate

TABLE 1--Continued

SUBORDINATE FACTORS

Instrument	Scale	Loading
	Factor VI	
RPM	Permissive (Super)	+.24
RPM	Authoritarian (Sub)	+.79
RPM	Equalitarian (Sub)	-.65
RPM	Permissive (Sub)	+.20
RPM	Ingratiator (Sub)	+.38
RPM	# of Trues (Sub)	+.92
RPM	# of Questions (Sub)	-.59
	Factor VII	
RPM	Ingratiator (Super)	-.22
SAS	Permissive (Super)	+.27
RPM	Critic (Sub)	-.27
RPM	Ingratiator (Sub)	-.20
RPM	# of Questions (Sub)	-.39
SRR	Authoritarian (Sub)	-.33
SRR	Equalitarian (Sub)	-.42
SRR	Permissive (Sub)	+.90
	Factor VIII	
RPM	Critic (Super)	-.22
RPM	Ingratiator (Super)	-.22
SBR	Rebel (Super)	+.23
RPM	Equalitarian (Sub)	-.21
RPM	Permissive (Sub)	-.32
RPM	Critic (Sub)	-.56
RPM	Rebel (Sub)	-.48
RPM	Ingratiator (Sub)	-.66
SRR	Authoritarian (Sub)	+.73
SRR	Equalitarian (Sub)	-.72
RI	Rebel (Sub)	+.91
RI	Critic (Sub)	-.72
RI	Ingratiator (Sub)	-.76
	Factor IX	
SRR	Permissive (Super)	+.22
SBR	Critic (Super)	+.86
SBR	Rebel (Super)	-.40
SBR	Ingratiator (Super)	-.32
RPM	Ingratiator (Sub)	-.23
	Factor X	
RPM	Authoritarian (Super)	+.20
SBR	Rebel (Super)	+.72
SBR	Ingratiator (Super)	-.80
RPM	Equalitarian (Sub)	-.43
RPM	Permissive (Sub)	+.21
RPM	Rebel (Sub)	+.52

TABLE 1--Continued

INTERACTION FACTORS

Instrument	Scale	Loading
	Factor XI	
RPM	Critic (Super)	+.31
RPM	Authoritarian (Super)	-.41
SRR	Permissive (Super)	+.85
SBR	Critic (Super)	+.24
RPM	Permissive (Sub)	+.35
RPM	Rebel (Sub)	-.22
RPM	Ingratiator (Sub)	+.24
SRR	Authoritarian (Sub)	+.43
SRR	Equalitarian (Sub)	-.33
	Factor XII	
RPM	Permissive (Super)	+.71
RPM	Critic (Super)	+.48
RPM	Rebel (Super)	-.30
RPM	Ingratiator (Super)	+.70
RPM	# of Trues	+.36
RPM	Authoritarian (Sub)	-.26
RPM	Critic (Sub)	+.26
RPM	Ingratiator (Sub)	+.23
SUB	# of Trues (Sub)	+.22
RPM	# of Questions (Sub)	-.36

DISCUSSION

This study failed to replicate some of the interactive effects of roles found in earlier research. Differences in the two squadrons which comprised the sample showed different interactional dynamics which tended to cancel each other out in several areas of expected relationships.

The results indicate at least two basic interactional patterns which are possible between each pair of superordinate-subordinate roles. The superordinate who occupies the authoritarian role and the subordinate who occupies the ingratiating role re-enforce each others needs in an adaptive way. This is the long term interaction which would be expected to emerge from the relationship. The authoritarian and rebel roles, however, are also positively correlated with each other because of a pro-active effort to counteract the role held by the other and would be short-term in value. Thus, any two roles of a reactive or adaptive component and a pro-active or counteractive component are in isolation and the analysis into these components will be the primary focus of future research.

REFERENCES

1. Cattell, R. B., Radcliff, J., and Sweney, A. B., Motivation components in children compared with those in adults. *Journal of General Psychology,* 1964, 70, 95-112.
2. Elsass, N., and Sweney, A. B., *Responsibility Index Handbook,* Test Systems Inc., 1972, Wichita, Kansas.
3. Elsass, N., and Sweney, A. B., *Supervise Ability Scale,* Test Systems, Inc., 1972, Wichita, Kansas.

4. Fiedler, F. E., *A theory of leadership effectiveness,* New York, McGraw-Hill, 1967.
5. Hemphill, J. K., Relations between the size of the group and the behavior of superior leaders. *Journal of Social Psychology.* 1950, 32, 11-22.
6. Lewin, K., Lippitt, R. and White, R., Patterns of aggressive behavior in experimentally created social climates. *Journal of Social Psychology.,* 1939, 10, 271-299.
7. McGregor, D. *Human side of enterprise,* New York, McGraw-Hill, 1960.
8. Stodgill, R. M. et al, *Aspects of leadership and organization,* Columbus: Ohio State University, 1953.
9. Sweney, A. B., Organizational power roles. *Professional Management Bulletin,* 10, 5-13, June, 1970.
10. Sweney, A. B., *Response to Power Measure Handbook,* Test Systems, Inc., 1971, Wichita, Kansas.
11. Sweney, A. B., *Supervisor Role Rating,* Test Systems, Inc., 1972, Wichita, Kansas.
12. Sweney, A. B., *Subordinate Behavior Rating,* Test Systems, Inc., 1972, Wichita, Kansas.
13. Sweney, A. B., Transactions of Power and Obligations, A paper presented to Mountain-Plains Management Conference, Boulder, Colorado, October, 1972.

NOTE

[1]Research reported here was sponsored by the Life Sciences Division of the Air Force Office of Scientific Research under Contract #2001.

Multivariate Experimental Clinical Research, Volume 2, Number 2, 1976, 101-106

MULTIVARIATE METHODOLOGY FOR N = 1

G. Frank Lawlis
The University of Texas Health Science Center at San Antonio

ABSTRACT

A multivariate design for N = 1 studies was presented for those areas of limited subjects or social psychological designs. The design called for a determination of expected outcomes to be used for baselines, and general criteria were related to the selection of control variables.

THE DESIGN

MULTIVARIATE METHODOLOGY FOR N = 1

As most introductory psychology students can readily describe, the goal of experimental design is to discover differences in behavior that are attributable to explainable or contrived phenomenon. A problem in many psychological research inquiries is that the researcher does not enjoy an unlimited supply of subjects who have a specified or defined characteristic. It would be a problem to order a sample of a strain of human beings with a given set of genetic traits much like one would order animals. Moreover, there are some populations that are finite with small numbers and no comparison normative data. For example, how many Patty Hearsts, or Richard Nixons, or Helen Kellers are there if a researcher had the resources and wanted to develop a research project around such an individual?

The case study approach is not new to psychology by any means, but the application of statistical application involves the whole issue of generalization of results. With this issue in mind the researcher has the challenge to eliminate extraneous variables that can compete for attribution in a change of behavior. The concept of control for N = 1 requires a variety of strategies maximizing sources of comparison. Three control references have been utilized in the past and are amendable to N = 1 studies; control or reference groups, pre-treatment baselines and conceptually equivalent control variables.

"A control group is a group that does not receive the experimental treatment, often providing a baseline from which the effect of the special treatment is measured." (Kimble *et. al.,* 1963, p. 35). The degree of sophistication in using control groups determines the degree in which the researcher is able to conclude changes in behavior is validity due to his experimental treatment. Campbell and Stanley (1963) summarize some strategies for effecting validity in studies; however, the more sophisticated the design becomes, the greater prerequisite for a large population for randomization of subjects. Ethical questions regarding the withholding services for research problems and isolation of certain people, as well as previously explained reasons of limited samples, make it difficult to employ such designs. It should be remembered that the utilization of a control group is for an expectancy probability of change. One method of estimating probability of change by chance alone for an individual would be to adopt the reliability coefficient determined by the normative sample of an instrument. Such a procedure requires a pre-investigation of stability.

Since reliability is generally reported as the correlation between trials of the same test, the prediction for any one score for the second trial with the knowledge of trial one would be the product of the correlation and trial one standard score.

$$Z_1 (r_{++}) = Z \text{ predicted } 2$$

The confidence range of r_{++} is found by converting to a Fishers Z_r (Ferguson, 1959, p. 151-152) and multiplying 1.96 times the standard error for the 95 percent confidence range.

$$Z_r \pm 1.96 S_{zr}$$

where

$$S_{zr} = \frac{1}{\sqrt{N - 3}}$$

By converting back to correlation coefficients the researcher could determine the limits of expected change scores for 95 percent confidence.

$$Z_1 (\pm r) = \text{range of expected outcome 95 percent of the time}$$

Note: The standard error for any individual score should also be taken into consideration.

Therefore, if the individual score of the second trial is outside the envelope of expectation, the researcher could make the conclusion that the subject's second score was beyond the expected probability of occurrence by chance.

The procedure of utilizing a normative group for comparison is one of utility; however, if one has the opportunity to observe and compute the individual's own predictability on past trials, the researcher can determine the likelihood of occurrence from history. Since we are considering an N of one, the probability of 95 percent confidence utilizes the number of observations as the sample of observations and determining the standard error from the individual's own behavior.

$$95\% \text{ Confidence range} = \bar{x} \pm 1.96 S\bar{x}$$

where

$$\bar{x} = \text{individual's average score over time}$$

$$Sx = \frac{\text{variance of observations}}{\sqrt{\text{number of trials}}}$$

Thirdly, the researcher could utilize a multivariate approach in which he could hypothesize and operationalize a set of experimental variables relevant to his treatment, and also operationalize a set of control variables that would serve as a baseline of behavior. The null hypothesis against which the research hypothesis would be paired would be if the change in behavior was due to history, maturation, testing, selection or mortality, then all variables would covary and change together.

The criteria for selecting control variables would be critical and depends upon the clarity with which the researcher can specify the treatment effects. Three criteria appear to be mandatory:

(1) There should be some evidence to support the assumption that, like control groups, the experimental and control variables are independent.

(2) The experimental variables must be conceptually equivalent and relevant to the treatment. There should be some reason to suspect that the control

variance could be affected if the treatment proved to be more generalized than predicted by the experimental variables alone.

(3) The experimental and control variables should be comparable to their perspective probabilities of change. That is, as opposed to the general problem of comparing one variable between two or more groups of randomized selection and having the assumption of common reliability, the underlying strategy being tested is that the experimental and control variable is equally sensitive to extraneous variance.

An extension of the N = 1 univariate analysis can be improvised. A z x z matrix of experimental and control levels with frequencies of outcomes discriminated into the expected range or beyond expected range. (See Figure 1).

Obviously, the table is amendable to a chi square analysis with expected outcomes as determined by the researchers. Such decisions would depend upon the number of variables and possibly the predetermined weightings of some variance relevant to the design.

An example of the application for the multivariate approach is the data taken from an evaluation of training program for one person. Ten experimental variables were selected as relevant and sensitive to the treatment, and ten control variables were selected for their relevancy to the program. The variable names and analysis are represented in Table 1.

FIGURE 1

	Expected Frequency	Beyond Expected
Experimental	Frequency	Frequency
Control	Frequency	Frequency

With the resultant analysis, the researcher could make the conclusion that the individual did change significantly, although he would be limited to extend the generalization beyond the subject.

NON-STABLE CONTROL VARIABLES

The application of control variance is not limited to those that are predicted to be stable from one observation to another. There are strong implications to situations in which growth functions can be compared. For example, the emotional maturation of adolescents may be compared to their physical growth. However, it would be necessary to transform the data matrix via analytic geometry to conform to a horizontal model. (See Figure 2).

TABLE 1

Training Effects on Individual Case

	Experimental Variable	Trial 1 (Z_1)	$r_{\#}$	Trial 2 (Z_2)
1	Empathy rating	-1.0	.90	1.21*
2	Warmth rating	0.0	.60	.62*
3	Genuineness rating	- .5	.90	.52*
4	Disclose rating	.5	.90	.45
6	Empath score	-1.0	.90	1.52*
7	Warmth score	1.0	.90	.58
8	Genuineness score	0.0	.90	1.05*
9	Disclosure score	-1.0	.80	.05*
10	Confrontation score	.5	.90	.08*
	Control Variable			
1	Outgoingness score	1.0	.85	.86
2	Elation mood	0.0	.60	.00
3	Seriousness score	- .5	.90	-.45
4	Inhibition score	-1.0	.85	-.85
5	Social Awareness	1.0	.90	.90
6	Tenseness score	.5	.80	.00*
7	Tension rating	1.0	.90	.00*
8	Tensions report	0.0	.80	-1.00*
9	Inhibition rating	- .5	.80	-.43
10	Inhibition report	0.0	.85	.00

*Beyond Expectation

Summary Table

	Expected	Beyond Expected
Experimental	2	8
Control	7	3

$x^2 = 5.05$, p2.05

Rotation of data matrices is done by first determining the number of degrees desired; in this case, the rotation needed to represent the control variable as linear horizontal data points. Since there will be two dimensions to be rotated, a 2 x 2 rotation matrix is required; however, an orthogonal relationship is to be maintained and the same degree of rotation will be assumed for both dimensions. The diagonal elements of the rotation matrix would be the cosines of the desired angles, whereas the rest of the elements would be the complement cosine values. If the rotation is counter clockwise (as in Figure 2), the upper right hand complement is given a negative sign, and if the rotation is clockwise the lower left hand complement is given a negative sign. Otherwise, the three remaining elements are positive. The resultant data matrix is simply post-multiplied by the rotation consine matrix (lamda).

	Data Values		Lamda		Rotated Values	
	x_1	y_1			x_i	y_i
Exp.	x_2	y_2	$\cos\theta°$	$\cos(90-\theta°)$		
	x_i	y_i				
	x_1	y_1	$\cos(90°-\theta°)$	$\cos\theta°$		
Control	x_2	y_2			x_j	y_j
	:	:				
	x_j	y_j				

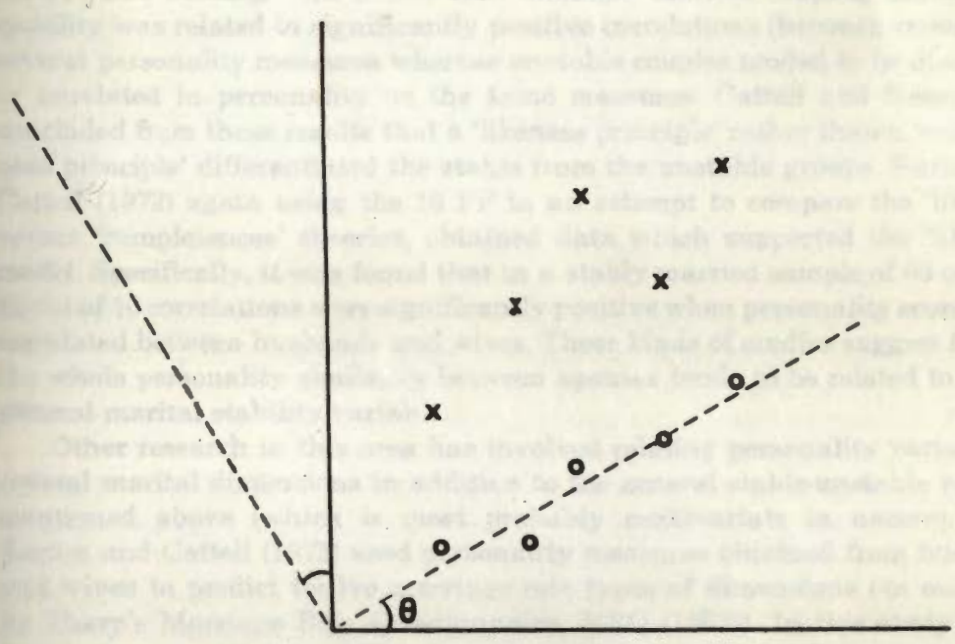
It should be pointed out that the resultant X or independent values will be different when compared to the original intervals of comparison. Therefore, in order to be able to compare variance at any one trial or the researcher is required to perform a linear transform per data point so as to analyze similar intervals of the independent variable. A crude transformation would be to apply the formula of proportions, such as:

$$\frac{\text{Rotated dependent value } (y')}{\text{Rotated independent value } (x')} = \frac{\text{Value of dependent value for prescribed independent value } (y'')}{\text{Desired interval measurement of independent value } (x'')}$$

or $\frac{y'}{x'} = \frac{y''}{x''}$

or $\frac{y'x''}{x'}$ = dependent value per determined value of independent value

FIGURE 2



ROTATION OF AXIS TO ASSUME HORIZONAL ANALOG

In summary, a multivariate design for N = 1 studies was presented for those areas of limited subjects or social psychological designs. The design called for a determination of expected outcomes to be used for baselines, and general criteria were related to the selection of control variable. It should be pointed out that such a design was not intended to be used as a substitute for more sophisticated designs. Control groups not only offer greater opportunities for generalization and sensitivity to experimental treatments, and in control variable designs, the researcher has to be more specific in his expected changes.

This paper should direct the implications to more extensive implications than merely to differences related to time events. The suggestion of a control variable has been advocated in behavior modification methodology, especially in individual case studies. The intended case for this presentation is when the researcher desires to serve two purposes: 1. To provide ad hoc hypotheses regarding a statistical analysis of time related data, and, 2. To provide a multivariate model for comparing various psychometric methods across time.

REFERENCES

1. Campbell, D. and Stanley, J. *Experimental and Quasi — Experimental Designs for Research.* Chicago: Rand McNally, 1963.
2. Ferguson, G. *Statistical Analysis in Psychology and Education.* New York: McGraw-Hill, 1959.
3. Kimble, G. and Garmegy, N. *Principles of General Psychology.* New York: Ronald Press, 1963.

Multivariate Experimental Clinical Research, Volume 2, Number 2, 1976, 107-111

PERSONALITY SIMILARITY IN SPOUSES RELATED TO MARRIAGE ROLES

Keith Barton
University of California, Davis

ABSTRACT

From an original sample of 93 married couples 71 couples were selected according to criteria which determined how similar they were to their spouses on several personality variables. Specifically 3 groups of people were defined: a Similar group who resembled their spouses highly in personality, a Random group who were neither positively or negatively matched with respect to personality measures and an Opposite group who tended to be opposite to their spouses on the personality measures. For each of these 3 groups, and for husbands and wives separately, scores on several marriage dimensions were compared. It was found that for both husbands and wives the Similar personality groups tended to score higher on the marriage role factor of Male Dominance than did either the Random or Opposite groups. For the wives data only, the Similar group also scored higher than the Random or Opposite groups on the marital factors of 1) Sexual Gratification and 2) Togetherness and Role Sharing. Possible implications and explanations for these findings are discussed.

INTRODUCTION

Considerable research has concerned itself with exploration of the relationships among various personality and marital variables. For example, Cattell and Nesselroade (1967) using Cattell's Sixteen Personality questionnaire (16 PF) and working with 'stable' and 'unstable' married couples, showed that stability was related to significantly positive correlations (between spouses) on several personality measures whereas unstable couples tended to be dissimilar or unrelated in personality on the same measures. Cattell and Nesselroade concluded from these results that a 'likeness principle' rather than a 'completeness principle' differentiated the stable from the unstable groups. Barton and Cattell (1972) again using the 16 PF in an attempt to compare the 'likeness' versus 'completeness' theories, obtained data which supported the 'likeness' model. Specifically, it was found that in a stably married sample of 69 couples, 12 out of 16 correlations were significantly positive when personality scores were correlated between husbands and wives. These kinds of studies suggest that on the whole personality similarity between spouses tends to be related to a very general marital stability variable.

Other research in this area has involved relating personality variables to several marital dimensions in addition to the general stable-unstable variable mentioned above (which is most probably multivariate in nature). Thus, Barton and Cattell (1973) used personality measures obtained from husbands and wives to predict twelve marriage role types of dimensions (as measured by Tharp's Marriage Role Questionnaire, MRQ (1963)). In this study it was found that most of the MRQ factors could be significantly predicted from a knowledge of the personality variables measured by the 16 PF. In the Barton and Cattell study no attempt was made to relate degree of similarity in

personality between spouses to marital dimensions. This present study was designed to take the Barton and Cattell data and reanalyze it, this time comparing the marriage role scores of groups of couples who differ in degree of personality similarity. In order to assign couples to groups which differ in their degree of personality similarity between spouses, use was made of a profile similarity coefficient (r_p) which allows one to identify couples who differ in the degree to which their personality *profiles* match, i.e. in this case, husbands and wives. Thus, a highly positive r_p, say .90, would indicate that the personality profile of the husband was very similar to that of the wife. On the other hand a negative r_p, say -.80 would indicate that the husband tended to have an opposite personality profile to the wife's. An excellent detailed description of this statistic is given by Tatsuoka (1974).

It is hypothesized that a group of married couples who exhibit a high positive profile similarity coefficient (r_p) on the 16 PF dimensions will differ from groups of married couples who have negative r_ps, or r_ps which differ little from zero. Specifically, the hypothesis is as follows: A highly similar group (between spouses on the 16 PF; *i.e.*, high positive r_p) will score higher than groups with lower similarity on the 16 PF on the following Marriage Role variables in the MRQ: Sexual Gratification, Togetherness and Role Sharing, Marital Stability, and Social Integration.

METHOD

SUBJECTS

One hundred eighty-six (93 couples) married graduate students from the University of Illinois were used as subjects. Each couple was paid ten dollars for participation in the study and was guaranteed anonymity of results. Only subjects who allowed their spouses to participate in the study were accepted.

TESTS

All subjects received the 16 PF and the MRQ tests, and both husband and wife completed each test together in the same session. A detailed report of the 16 PF factors can be found in the handbook (Cattell, Eber, and Tatsuoka, 1970). The MRQ factors are outlined in Table 2 but the reader is directed to the original Tharp study (1963) or the Barton, Cattell, and Kawash (1971) article for a more detailed report.

PROCEDURE

A profile similarity coefficient (r_p) was calculated for each of the 93 couples. On the basis of this r_p couples were assigned to one of the following groups:

1. A Similar Personality group ($r_p \geq .241$). N = 32
2. A Random Personality group ($.10 > r_p > -.066$). N = 25
3. An Opposite Personality group ($p_{rp} \leq -.066$). N = 14

The cut off points for the groups were selected after visual inspection of the distribution of r_ps, and resulted in removing from the analysis 22 couples who met none of the three group inclusion criteria. It was felt that it was more important to make the groups as different as possible in terms of r_p rather than to include all the couples. One way analyses of variance were calculated with three levels of personality similarity as the factor in the design (independent variable) and the 12 MRQ factors as dependent variables. Separate analyses were made for the husbands and wifes' data.

Table 1

MRQ Factors: A Brief Description

1. Sexual Gratification
   Loads on items indicating fulfillment in sex and affection.

2. Togetherness and Role Sharing
   Loads on items indicating agreement between spouses in such adverse areas as finance, recreation, religion, friends, children, manners, general philosophy, etc.

3. Home Devotion
   Loads on items indicating that the wife has a high interest in keeping the home clean and neat and that she does most of the housework.

4. Participation in Community Affairs
   Items indicate much social interaction with the community, and friends.

5. Social-Intellectual Equality
   Items indicate perceived equality between spouses on IQ and social strata.

6. Marriage Instability
   Items indicate that separation and/or divorce have been contemplated. Other item reflects visits to marriage counselors and a statement of general satisfaction with the marriage.

7. Social integration
   Items indicate that all members of the family (including the children) do things together often.

8. Work Performance
   Items reveal whether or not each member of the family has his own set jobs to do.

9. Social Influence
   Items show which spouse has the most social influence on the other, i.e. who influences most the choice of friends, etc.

10. Spouse Independence
    Items show who controls the money, number of children planned, etc.

11. Wife Adequacy
    Items refer to the way in which wives participate in such matters as who earns income, etc.

12. Male Dominance
    Items reveal which spouse is the main influence in such matters as spending, sex, childrearing, et.

RESULTS

The results of the analyses of variance are shown in Table 2. As can be seen from Table 2 for both husbands and wives the groups differed on the MRQ scale of Male Dominance. In addition, for wives only, the groups differed on MRQ factors 1 and 2 (Sexual Gratification and Togetherness and Role Sharing). In order to explore which groups were responsible for the overall F ratios being significant, comparisons of all parts of means (on each of the MRQ factors) were

made using the conservative Tukey (a) procedure. In the case of the Male Dominance factor, these analyses revealed that for both husbands and wives the Similar groups scored higher than the Random groups ($p < .01$ in both cases). For wives only, the Similar group also scored higher than the Opposite group ($p < .05$).

In the case of the Sexual Gratification factor (for wives only) the Similar group scored higher than either the Random or Opposite groups ($p < .01$ in both cases). For the Togetherness and Role Sharing factor again the Similar group scored higher than both the Random and Opposite groups ($p < .05$ in both cases; wives data only).

Table 2

Analyses of Variance for Husbands' and Wives' Data

Husbands' Data

	Source	DF	SS	MS	F	p
1. MRQ factor 12						
Male Dominance	Between Groups	2	6.00	3.00	4.58	0.013
	Within Groups	68	44.55	0.65		
	Total	70	50.55			

Wives' Data

	Source	DF	SS	MS	F	p
1. MRQ factor 1						
Sexual Gratification	Between Groups	2	7.79	3.90	3.59	0.032
	Within Groups	68	73.92	1.09		
	Total	70	81.71			
2. MRQ factor 2	Source	DF	SS	MS	F	p
Togetherness and	Between Groups	2	6.38	3.19	3.96	0.02
Role Sharing	Within Groups	68	54.80	.81		
	Total	70	61.18			
3. MRQ factor 12	Souces	DF	SS	MS	F	p
Male Dominance	Between Groups	2	7.02	3.51	4.63	0.013
	Within Groups	68	51.61	.76		
	Total	70	58.63			

MANUSCRIPT INFORMATION

Submission and Review. Manuscripts should be prepared in triplicate following the instructions contained in the most recent edition of the *Publication Manual of the American Psychological Association.* Manuscripts will be forwarded for reading by two members of the Board of Consulting Editors selected for their expertise in the area covered by the paper. Editorial decisions should be received by the author(s) within six weeks of submission.

Length. Although papers should be as concise as clarity permits, and tables, figures, and graphs kept as few as possible, there is no standard limit on length. However, manuscripts less than 1500 words with no more than two tables, figures, or graphs may receive earlier publication.

Abstract. All manuscripts should be preceeded by an abstract not to exceed 150 words.

Reprints. Fifty reprints of each manuscript accepted for publication are provided free of charge. Additional reprints in quantities of fifty may be ordered on the form sent out with proofs provided the order is returned within seven days of receipt of proofs.

Publication Charge. Authors will pay, prior to publication, a part of publication costs assessed currently at $15 (U.S.) per manuscript page plus $15 (U.S.) per table, figure, or graph. These charges are subject to change without notice.

Tables, Figures, and Graphs. Tables, figures, and graphs should, except in rare cases, be camera ready.

General. Submission of any manuscript will be taken to imply that it is unpublished and is not being considered for publication elsewhere. Papers published in *Multivariate Experimental Clinical Research* may not be reprinted or published in translation without permission from the Editor. This restriction is waived for authors who wish to reproduce their own articles and when material is reproduced in limited quantity for instructional purposes. Contributions are welcomed from researchers of all nationalities, but must be written in English.

SUBSCRIPTIONS

Subscriptions are available on a per volume basis. The subscription rates are $16 (U.S.) per volume for subscribers in the United States and $22 (U.S.) for subscribers outside of the United States. Single issue price when available is $4 (U.S.) for the U.S. and $5.50 (U.S.) for out of U.S. Prices are subject to change without notice.

ADDRESS FOR MANUSCRIPT SUBMISSION AND SUBSCRIPTIONS

Dr. Charles Burdsal, Editor
Multivariate Experimental Clinical Research
Department of Psychology, #34
Wichita State University
Wichita, Kansas 67208
U.S.A.

TABLE OF CONTENTS