ERROR CORRECTION

Guilford, J. P. *A Sixty-Year Perspective on Psychological Measurement,* Volume 9, Number 4, pp. 341–349.

Due to a typesetting error, several pages in this article are out of order.

The pages should be read in the following order:

341, 344–345, 342–343, 346–349.

Readers should remove this page and insert it in their copy for future reference.

A Sixty-Year Perspective on Psychological Measurement

J. P. Guilford
University of Southern California, Emeritus

Editor's Note. This is the second article in our Perspective on Psychological Measurement Series. In this article J. P. Guilford, who was first exposed to the problems and methods of psychological measurement in the early 1920s, provides a unique perspective on his distinguished career ranging across his varied and very important contributions to the science of psychology and to the methodology of psychological measurement.

Based upon experiences with most kinds of methods of psychological measurement, this article presents comments on a variety of uses, including psychophysics, scaling, testing, and factor analysis. Some difficulties are pointed out, some faults are mentioned, and a variety of applications are discussed, some of them unusual.

With the reader's indulgence, I will begin with some personal notes. My first awareness of the possibility of psychological measurement came in a course on psychophysics under Karl M. Dallenbach at Cornell University. This experience revealed the possibility of assigning numbers to psychological events with high degrees of precision.

In Dallenbach's course on memory another determining experience occurred. He suggested that memory span could be conceived as a limen and could be measured by application of the method of constant stimuli. I proceeded to do the needed experiment. I read to a number of student subjects lists of digits of varied lengths in order to find for each subject the proportion of correct responses at each series length. The proportions did follow the cumulative normal distribution, from which a median and a standard error could be obtained (Guilford & Dallenbach, 1925).

Preceding the Cornell events I had had some extensive experiences with psychological tests at the University of Nebraska. The testing of student aptitudes had just begun about 1920. The psychology staff was involved in administering the Army Alpha Examination and Thorndike's tests for college students. As an assistant, I was involved in scoring and in dealing with the results. During two years as a graduate student, believe it or not, I was in charge of the Psychological Clinic and administered individual intelligence tests to children.

Experiences from the two quite distinct directions eventuated in my writing *Psychometric Methods* in 1936. Coincidentally, this was the time things were generally stirring in psychological measurement, as indicated by the founding of the Psychometric Society and the journal *Psychometrika* at about the same time.

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 9, No. 4, December 1985, pp. 341–349

0146-6216/85/040341-09$1.70

tests of known factorial composition. Success for the assignment to pilot training, for example, yielded multiple correlations approaching .60. As a general result of the program, the rate of failure in pilot training, which had been about 30% before the war, was lowered to about 10% before the end of the war.

Similar results have been obtained more recently on the general educational scene, a few of which will be cited. In 1965 (Guilford, Hoepfner, & Peterson, 1965), we found multiple correlations approaching .60 for predicting ninth grade mathematics test performance. Hills (1957) found many significant correlations in the prediction of mathematical achievement at the senior and graduate level in college. Caldwell, Schraeder, Michael, and Meyers (1970) found multiple correlations of .60 in each of two schools for predicting achievement in geometry, using five structure-of-intellect (SOI) scores in one case and four in the other. Ignatz (1982) found that prediction of achievement in physics from two hours of SOI testing was as effective as that from *two days* of testing with other tests. For a somewhat different use of SOI factors, Wiig and Semel (1976) found the concepts very useful in understanding speech disorders of children. I have maintained that the use of SOI factors provides psychology in general with unambiguous operational concepts (Guilford, 1982).

Factor Analysis

I doubt that anyone has had more experience in performing factor analyses than I. In my aviation-psychology experiences during World War II, there were 15 large analyses of human abilities as related by Guilford and Lacey (1947) and by French (1951). I have commented elsewhere (Guilford, 1948a) on the great value of using factor analysis in a test-development program for the selection and classification of personnel. One of the great advantages is that in the prediction of achievement, scores for the different contributors have minimal intercorrelations, and the sources of validity are meaningful variables. With the availability of the SOI model, needed abilities and as yet undemonstrated abilities can be readily hypothesized.

Later, in the Aptitudes Research Project (ARP) at the University of Southern California, which lasted for 20 years, there were some 40 analyses of similar scope. There have been others, for example, an analysis of human interest variables (Guilford, Christensen, Bond, & Sutton, 1954) and of temperament (Guilford & Zimmerman, 1956).

From time to time I have spoken out with regard to what I considered to be faults and misuses of factor analysis (Guilford, 1952; and on other occasions to be mentioned). One of my chief concerns is that factor analysis should be employed as an experimental investigation, with appropriate controls in the population of subjects to be examined, and with respect to test construction for the purposes of the analysis. Building controls into the tests to be used is an exacting process. The latter should proceed from hypotheses set up in planning the analysis. That is, the kinds of factorial variables should be envisioned, with alternative conceptions as to the nature of each expected factor. Too many analyses have been performed as afterthoughts or just because some intercorrelations were available.

I have differed from most of the others who factor analyze in the use of orthogonal rotations of axes and in foregoing computerized procedures. Orthogonal rotations have not been done with the assumption that factor axes should be at right angles, the factors being entirely independent. Experience has shown that when an appropriate combination of tests has been utilized, there are usually rather wide separations between clusters of tests, and that orthogonal axes offer good approximations to the factors represented by the sets of tests. There is also a lack of trust in locating oblique axes where each one is represented by a pitifully small sample of test vectors. I have been comforted by the fact that Thurstone found his original set of primary abilities by means of orthogonal, graphic rotations and they have mostly stood the test of time. In general, I have found good replications of factors in successive analyses in relation to their representative tests.

Computerized Rotations

The story of my rejection of computerized rotations is not a short one. Before such rotations became available, we always used Thurstone's (1927) graphic method, modernized by the device that Zimmerman (1946) contributed to facilitate the procedure. As computerized methods became more generally available during the ARP, we tried most of them. None proved to be satisfactory for interpretability of factors or for replications of factors. I will relate below specific experiences to support my case.

Concerning computerized rotations in general, with Zimmerman (Guilford & Zimmerman, 1963), I demonstrated that the location of rotated axes is very much dependent upon the peculiar sampling of tests put into the test battery. More recently (Guilford, 1977a), I did a minor study with the use of 13 small contrived factor matrices. Intercorrelations were computed, to which small realistic random errors were added. The number of univocal "tests" was varied from none in four matrices to a few in others. Both Varimax and Promax solutions were applied, the latter to represent oblique rotations.

The findings included the fact that only when all tests in an analysis were univocal was there good reproduction of the original matrices. In some analyses, some factors were badly distorted and some were entirely lost. An incidental finding was that the scree test of when to stop extracting factors often underestimated the correct number of factors. In my experience it has appeared that in order to achieve the best results, it is better to extract a larger number of factors than the scree test calls for; better too many than too few, lest a meaningful factor be lost. Any uninterpretable rotated factor can be rejected. Meanwhile, one has made use of all contributing information.

Toward the end of the analyses by the ARP, Cliff's (1966) method of rotating by computer to hypothesized factor matrices became available. Our hypotheses were naturally derived from the SOI model, since the tests had been designed for SOI abilities. After each rotation of the axes has been made, inspection of the results indicates where both meaning and the appearance of simple structure might be improved. This process is repeated until a satisfactory fit has been achieved. In these respects the process is very much like graphic rotations. We rerotated all principal-axes matrices previously rotated by the graphic method and found excellent agreement.

In 1969, (Guilford & Hoepfner, 1969), I compared the results of rotating by the Cliff (1966) method with those from Varimax (Kaiser, 1958) solutions in several analyses. The comparisons were very illuminating. A general finding was that Varimax tended to overdo the demands for simple structure. That is, it obtained too many zero loadings and threw the large loadings toward a small number of factors. With respect to the production of positive manifold, which should be the case with abilities, it often ended with too many unacceptable negative loadings. In 12 analyses there were 14 negative loadings more negative than −.30.

As an example of the overshooting of simple structure, in one analysis with 59 tests and 25 extracted factors rotated by Varimax, the results included two factors with 17 significant (greater than +.30) loadings each, six factors had only two significant loadings, and five factors had none.

As to interpretations of rotated factors, in general the Varimax method showed a very poor record. Overall, only 32% of the factors could be interpreted as SOI abilities, whereas with the Cliff (1966) method the batting average was 93%. To cite some specific instances, it is well recognized that a vocabulary test should indicate probably the most generally recognized factor of verbal comprehension, or cognition of semantic units (CMU), in the SOI model. In 20 analyses this test came out in only 7 cases by Varimax rotations.

To cite some further instances, the Plot Titles test, scored for the number of nonclever responses, commonly helped to determine the SOI ability divergent production of semantic units (DMU), but it came out on that factor in only 7 of 12 cases. The count of the number of clever responses is customarily a score for divergent production of semantic transformations (DMT), but it came out as such in only 1 of 12 instances by the Varimax method.

The use of the Cliff (1966) method of rotations has encountered severe criticism from Horn and Knapp (1973) because of the "subjective element" involved. For one of our matrices of extracted fac-

In what follows I shall take some stocktaking regarding methods of psychological measurement within the limits of my own experiences with those methods. I shall make some references to psychophysics, scaling methods, tests, and factor analysis. The methods were employed in connection with both basic science and psychotechnology. I examined the early years in psychometric development about 25 years ago (Guilford, 1961), with some suggestions for future directions. The present article should bring my views up-to-date.

Psychophysics

My flair with psychophysics ended essentially at the time of the publication of the second edition of *Psychometric Methods* in 1954 (Guilford, 1954a). In that edition I dealt with the need for revisions in both Weber's and Fechner's laws. I stated that the "just noticeable difference," or difference limen, indicated by ΔS, is not in a simple ratio to the starting point S, but is better expressed by a law that states that $\Delta S = S^n$, where n is expected to vary between 0 and 1, which had been commonly found experimentally. Integrating this equation gave $R = aS^b + c$, where R stands for the sensory response. This expression was proposed as a replacement for Fechner's law.

With Dingman (Guilford & Dingman, 1954), I tested the validity of this power function by applying the method of absolute judgment to lifted weights. The law was fully supported. Later, Stevens and his associates (e.g., 1957) also demonstrated the law's validity within various sensory areas. To my knowledge there has been no reason for revision of these two laws. They are quite general, of course. Previous laws had held to an exponent equal to 1.0. It might have been of interest to explore reasons for differences in that value.

In my dealings with psychophysical theory I found it very useful to think in terms of Thurstone's (1927) law of comparative judgment. This has been especially true in thinking of scaling methods in psychology. An example will be mentioned below. In view of Thurstone's numerous contributions to psychological measurement I have often thought that there should have been more psychologists who were trained in engineering first.

Psychological Scaling Methods

Psychological scaling methods, such as ratings, equal-appearing intervals, and ratio judgments, attempt to calibrate psychological events as estimated by observers. Assigning numbers to such information involves careful operations and consideration of the actual kinds of scales resulting from such efforts. It is well to keep in mind the types of scales distinguished by Stevens (1946), whether the numerical values obtained pertain to nominal, ordinal, interval, or ratio scales. The type attained determines the kinds of mathematical treatment that are justified and the kinds of conclusions that can be drawn.

Rating Scales

Rating scales are evidently the most common method of measurement employing human judgments of quantity. Under the best of circumstances, such values may be accepted as being on an interval scale, but constant vigilance is needed to ensure that status, and the variables on which measurements are desired must not be taken for granted.

One of the most common technical uses of ratings is aimed at evaluating personnel with respect to traits or performances of some kind. For example, in one study the traits were eight defined thinking abilities that had been well distinguished by factor analysis (Guilford, Christensen, Taaffe, & Wilson, 1962). Seven supervisors of research scientists were asked to rate the scientists under their supervision as to standings on those traits. The common result for any supervisor was that he/she gave ratings with high correlations among the traits, indicating strong halo effects. There may have been some genuine correlation among the thinking traits, but certainly not that much.

Ratings in psychodynamics. My experiences with rating methods have been mostly in the area of research that I have called "psychodynamics" (Guilford, 1939). This area of psychological functioning is concerned with the affective values or preferences for sensory experiences of different characteristics. My investigations have been concerned with kinds of colors and color combinations and with odors and odor combinations.

At times there have been some difficulties that should be mentioned in connection with affective scales. On such scales there is commonly a middle category labeled ''indifferent,'' indicating neither pleasant nor unpleasant. This has been known to lead to strongly bimodal distributions of judgments, in avoidance of the indifferent category (Guilford, 1938). Reminding raters to avoid rejection of this category or computing medians should solve the problem.

Being concerned with discovering functional relationships, I wished to relate affective value to specified aspects of stimuli or to measures of psychological aspects of sensations. Colors have long been recognized as existing in a three-dimensional system, with graded variations in each aspect, as in the Titchenerian color pyramid. A quantitative treatment of this model was provided by the *Munsell Book of Color* (1929). This source provides samples of 20 psychologically equal steps as to hue, in the color circle extending from red around to purple-red. The scale for brightness or lightness extends in a line from white to black. Saturations or chromas for the various hues, at different graded color-intensity levels, appear at every hue.

For me, the *Munsell Book of Color* (1929) served as a challenge. Instead of being able just to make very rough statements, such as that preferred colors are red, green, and blue, and less liked colors are yellow, blue-green, and violet, it should be possible to determine whether there are any continuous functions in this area of sensory functioning.

A large number of samples of colors were chosen at alternate hues and at chosen lightness and chroma levels. The color squares on a gray background were exposed to observers for their ratings on a 9-point scale. Results from these experiments have been reported in a number of places (Guilford, 1934, 1940b, 1949; Guilford & Smith, 1959). Continuous functions were found for the relation of affective values to hue, lightness, and chroma.

The relations to hue, at the middle lightness and chroma levels, were obviously periodic in nature and could be fitted to Fourier functions with the first and third components involved. Curves relating affective value to lightness and chroma were of second-order types. At the 10 hue steps isohedon charts were produced, showing the affective values at constant levels for various levels of lightness and chroma. These were drawn for men and women separately, although gender differences were small. There has been interest in these findings especially for motion-picture engineers and others concerned with optics (Guilford, 1940a).

An example of a similar application of measurement can be cited from the field of audition. Singer and Young (1941) obtained a large number of judgments of the affective values of single tones varying widely in pitch and in loudness. Although their judgments were only in terms of liking or disliking, by applying Thurstone's law of comparative judgment I converted the proportions of liking to values on an assumed interval scale. With smoothing of the trends in the values, curves relating affective value to pitch and to loudness could readily be obtained. Isohedon charts showing loci of equal levels of preference were constructed, like those for color preferences (Guilford, 1954b).

The reporting of these studies is to emphasize the fact that many such problems, in experimental psychology, especially, have been overlooked, for apparent lack of interest. In the case of the sound-preference problem, for example, we can conclude something more precisely than merely to say that low-pitched and soft sounds are preferred.

Psychological Tests

To appraise the numerous contributions to psychological test development and use would require a book-length document. I can touch on only a few points on this subject. In many places I have strongly urged that with tests we should aim as much as possible at the measurement of psychological variables determined by factor analysis (e.g., Guilford, 1940b, 1948b). More recently, I have held that ordinary nonfactorial variables tend to be ambiguous and are thus poor elements for good communication (Guilford, 1982).

For example, I have cited experiences in applying tests for the selection and assignment of students for training in the U.S. Army Air Force during World War II (Guilford & Lacey, 1947). The practice was to determine aptitude scores for training as pilots, navigators, bombardiers, and others in terms of weighted composites of scores from tors, they generated a random factor matrix toward which the computer rotated. They claimed a good fit of obtained loadings to targeted loadings. I have more recently responded to their criticism (Guilford, 1981b). I could fortunately cite a report from Elshout, Van Hemert, and Van Hemert (1975), in which they objected that Horn and Knapp had used only one randomly-generated target matrix. Having a computer generate a large number of targeted matrices, and determining an index of goodness-of-fit for each rotation, they found that the Horn-Knapp solution gave an index far outside the whole range of indices that they obtained.

g Factors

In the literature there are still indications of some belief in Spearman's (1904) universal *g* ability, or "general intelligence." I do not see how anyone who has administered individual tests to children, as I did as a graduate student, could believe in such a unitary ability. Nor could anyone who knows the meaning of correlations and who knows about the large proportion of zero correlations among tests of an intellectual nature believe in a *g* factor.

As a simple test of whether a *g* ability could be detected by factor analysis, in 1941 (Guilford, 1941) I produced an arbitrary small matrix that contained a completely general factor. The latter was given to the students in my class on factor analysis as an exercise. Every student reported finding a general factor. Thus, *g* could be evident at the basic or first-order level, yet no one that I know of has ever reported such an instance with real data.

In more recent years it has been a common idea that a *g* factor needs to be demonstrated as a higher-order factor. In my own experience with higher-order factors (Guilford, 1981a), the most general higher-order factor that I encountered was at the third-order level. There were many indications that in my SOI model there are 16 third-order factors: one for each category of abilities, for five operations, five kinds of informational content, and six kinds of informational products. There is some suggestion of some degree of correlation between the operation factors of cognition and evaluation. However, such a relationship could be attributed to the fact that those kinds of abilities are commonly measured by multiple-choice test formats. For the other operation categories, other test formats are more common.

The best evidence against the belief in a genuine *g* factor is a substantial proportion of zero correlations among tests of intellectual abilities. This was most recently pointed out by Hoepfner and me (Guilford & Hoepfner, 1971). From experiences in the ARP, there were more than 48,000 coefficients of correlation involved. Of these, about 18% lay between −.10 and +.10. Many more failed to reach the .05 level of significance, and could thus be accepted as zero. It does not seem possible to believe in a completely general *g* ability.

And how about *g* factors of lesser scope, as proposed by Horn and Cattell (1966)? They have championed two less general factors, g_c (crystallized intelligence) and g_f (fluid intelligence). The former is said to be determined by experience, including education, and the latter more determined by heredity. Verbal tests are typical of the former and nonverbal tests of the latter. Incidentally, they obtained these factors by testing individuals who varied in age from adolescents to senior citizens. All these facts gave me the clue to suggest that their g_c factor is equivalent to my third-order factor of general semantic ability and that their g_f is a confounding of my third-order factors for visual and symbolic information (Guilford, 1980). There is also the fact that semantic abilities tend to hold up better with advancing age.

Morphological Models

I should like to emphasize the great advantages of a strategy that I have found so useful in a factor analysis research program. This is, of course, my SOI model (Guilford, 1967). As should be generally known by now, the model is three-dimensional, one facet for kinds of mental operations, one for kinds of informational content, and the third for kinds of informational forms or products. I have also found such a model useful with psychomotor abilities (Guilford, 1958), with kinds of temperament factors and types of psychopathology (Guilford, 1959b).

The heuristic value of the SOI model in predicting as yet undemonstrated factorial abilities has been amply demonstrated. In all the years preceding announcement of the model in 1959, some 40 such abilities had been demonstrated. Because the model pointed clearly toward additional abilities, within the next 10 years at least 60 had been added to the list. Thus, with 150 abilities projected by the model (Guilford, 1977b), some 50 are calling for investigation. It is possible that the number will have to be extended in order to include a possible 30 kinesthetic abilities; a rich field for future research.

The SOI form of model qualifies as a mathematical "product of sets." The periodic table of elements in chemistry is another example of it. It has been recommended by the astronomer Zwicky (1957), who used it in his science. It has been proposed as a useful aid in creative thinking, by Allen (1962) and others. I can well vouch for its utility, as related above.

Q Factor Analysis

As compared with the more commonly known *R* analysis, which uses the correlations between tests, *Q* analysis, which is based on correlations between individuals, rarely occurs. Yet there are useful places for it. This was illustrated by such an analysis (Guilford & Holley, 1949) dealing with aesthetic preferences. A number of men and women were asked to rate their likes and dislikes for designs on the backs of playing cards. Correlations were obtained among the raters and *Q* factor analyses were performed. Some of the factors could be interpreted as different *themes* among the designs. A possible application of such an analysis could be to preferences in the marketplace.

A quite different application has been reported concerning stereotypes of individuals in different occupations—insurance salespeople, clergy, physicists, engineers, and journalists. Seven individuals from each occupation, who were well established in their lines of work, were selected. Each person answered the items in the Strong Vocational Interest Blank and also some items from temperament scales (J. S. Guilford, 1967). The *G* index of agreement (Holley & Guilford, 1964) was used in correlating the individuals. This index of correlation was used instead of the phi coefficient because it avoids the biasing effects of differences between means. The five rotated factors were very much in line with the five occupational groups. An examination of the kinds of responses favored by each occupational group revealed the nature of the occupational stereotype.

After the *G* index had been proposed, Holley made numerous applications of it in analysis of groups of individuals in different psychopathological categories. Holley and I (1966) have elaborated on the rationale for *G*. The results of analyses with pathological groups are in the form of syndromes of symptoms. One interesting finding by this approach was a new kind of evidence for validity of the Rorschach test (Holley, Frobarj, & Ekberg, 1965).

Concluding Remarks

Psychological measurements have gained considerably in maturity and in spheres of usefulness in a great variety of places, in both basic science and in psychotechnology. Much vigilance is still needed, however, in order to avoid miscarriages and to promote the usefulness of psychological measurements. Full advantage has not been taken of the methods, particularly in our basic science. It is to be hoped that these special articles in this Journal will do much to steer the uses of the various methods and to encourage their applications.

References

Allen, W. S. (1962). *Morphological creativity*. Englewood Cliffs NJ: Prentice Hall.

Caldwell, J. R., Schraeder, D. R., Michael, W. B., & Meyers, C. E. (1970). Structure-of-intellect measures and other tests as predictors of success in tenth-grade modern geometry. *Educational and Psychological Measurement, 30,* 437–441.

Cliff, N. (1966). Orthogonal rotation to congruence. *Psychometrika, 31,* 33–42.

Elshout, J. J., Van Hemert, N. A., & Van Hemert, M. (1975). Comment on Horn and Knapp: On the empirical base of Guilford's structure-of-intellectual model. *Tjidschrift voor Onderwijsresearch, 1,* 15–25.

French, J. W. (1951). The description of aptitude and

achievement tests in terms of rotated factors. *Psychometric Monographs*, No. 5.
Guilford, J. P. (1934). The affective value of colors. *Journal of Experimental Psychology, 17*, 342–370.
Guilford, J. P. (1936). *Psychometric methods*. New York: McGraw-Hill.
Guilford, J. P. (1938). The computation of psychological values from judgments in absolute categories. *Journal of Experimental Psychology, 22*, 32–42.
Guilford, J. P. (1939). A study in psychodynamics. *Psychometrika, 4*, 1–23.
Guilford, J. P. (1940a). Human abilities. *Psychological Review, 47*, 367–394.
Guilford, J. P. (1940b). There is system in color preferences. *Journal of the Optical Society of America, 30*, 455–459.
Guilford, J. P. (1941). A note on the discovery of a *g* factor by Thurstone's centroid method of factor analysis. *Psychometrika, 6*, 205–208.
Guilford, J. P. (1948a). Factor analysis in a test-development program. *Psychological Review, 55*, 79–84.
Guilford, J. P. (1948b). Some lessons from aviation psychology. *American Psychologist, 3*, 3–11.
Guilford, J. P. (1949). System in color preferences. *Journal of the Society of Motion Picture Engineers, 52*, 197–210.
Guilford, J. P. (1952). When not to factoranalyze. *Psychological Bulletin, 49*, 26–37.
Guilford, J. P. (1954a). *Psychometric methods* (2nd ed.). New York: McGraw-Hill.
Guilford, J. P. (1954b). System in the relationship of affective value to frequency and intensity of auditory stimuli. *American Journal of Psychology, 67*, 691–695.
Guilford, J. P. (1958). A model for psychomotor abilities. *American Journal of Psychology, 71*, 164–174.
Guilford, J. P. (1959a). *Personality*. New York: McGraw-Hill.
Guilford, J. P. (1959b). Three faces of intellect. *American Psychologist, 14*, 469–479.
Guilford, J. P. (1961). Psychological measurement a hundred and twenty five years later. *Psychometrika, 26*, 109–127.
Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
Guilford, J. P. (1977a). The invariance problem in factor analysis. *Educational and Psychological Measurement, 37*, 11–19.
Guilford, J. P. (1977b). *Way beyond the IQ: Guide to improvement of intelligence and creativity*. Buffalo: Creative Education Foundation.
Guilford, J. P. (1980). Fluid and crystallized intelligence: Two fanciful concepts. *Psychological Bulletin, 88*, 406–412.
Guilford, J. P. (1981a). Higher-order structure-of-intellect abilities. *Multivariate Behavioral Research, 16*, 411–435.
Guilford, J. P. (1981b). A map for an intelligent education. *Education, 101*, 310–314.
Guilford, J. P. (1982). Cognitive psychology's ambiguities: Some suggested remedies. *Psychological Review, 89*, 48–59.
Guilford, J. P., Christensen, P. R., Bond, N. A., & Sutton, M. A. (1954). A factor analysis of human interests. *Psychological Monographs, 68* (Whole No. 375).
Guilford, J. P., Christensen, P. R., Taaffe, G., & Wilson, R. C. (1962). Ratings should be scrutinized. *Educational and Psychological Measurement, 22*, 439–447.
Guilford, J. P., & Dallenbach, K. M. (1925). Determination of the memory span by the method of constant stimuli. *American Journal of Psychology, 36*, 621–628.
Guilford, J. P., & Dingman, H. F. (1954). A validation study of the ratio-judgment method. *American Journal of Psychology, 67*, 395–410.
Guilford, J. P., & Hoepfner, R. (1969). Comparisons of varimax rotations with rotations to theoretical targets. *Educational and Psychological Measurement, 29*, 3–22.
Guilford, J. P., & Hoepfner, R. (1971). *The analysis of intelligence*. New York: McGraw-Hill.
Guilford, J. P., Hoepfner, R., & Peterson, H. (1965). Predicting achievement in mathematics from measures of intellectual-aptitude factors. *Educational and Psychological Measurement, 25*, 659–682.
Guilford, J. P., & Holley, J. W. (1949). A factorial approach to the analysis of variance in esthetic judgments. *Journal of Experimental Psychology, 38*, 208–218.
Guilford, J. P., & Lacey, J. I. (1947). *Printed classification tests: Army Air Force Aviation Psychology Research Program* (Report No. 5). Washington DC: Government Printing Office.
Guilford, J. P., & Smith, P. C. (1959). A system of color preferences. *American Journal of Psychology, 72*, 487–502.
Guilford, J. P., & Zimmerman, W. S. (1956). Fourteen dimensions of temperament. *Psychological Monographs, 70* (Whole No. 417).
Guilford, J. P., & Zimmerman, W. S. (1963). Some variable-sampling problems in the rotation of axes in factor analysis. *Psychological Bulletin, 60*, 281–301.
Guilford, J. S. (1967). Isolation and description of occupational stereotypes. *Occupational Psychology, 41*, 57–64.
Hills, J. R. (1957). Factorized abilities and success in college mathematics. *Educational and Psychological Measurement, 17*, 615–622.
Holley, J. W., Frobarj, G., & Ekberg, K. (1965). On

the validity of the Rorschach test. *Scandanavian Journal of Psychology, 6,* 7–18.

Holley, J. W., & Guilford, J. P. (1964). A note on the G index of agreement. *Educational and Psychological Measurement, 24,* 749–753.

Holley, J. W., & Guilford, J. P. (1966). Note on the double centering of dichotomized matrices. *Scandanavian Journal of Psychology, 7,* 97–101.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology, 57,* 253–270.

Horn, J. L., & Knapp, J. R. (1973). On the subjective character of the empirical base of Guilford's structure-of-intellect model. *Psychological Bulletin, 80,* 33–43.

Ignatz, M. (1982). A comparison of S-O-I factors with the Florida Statewide Twelfth Grade Test in predicting physics achievement. *Educational and Psychological Measurement, 42,* 935–939.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation of axes in factor analysis. *Psychometrika, 23,* 187–200.

Munsell Book of Color. (1929). Baltimore: Author.

Singer, W. B., & Young, P. T. (1941). Studies in affective reaction II. *Journal of General Psychology, 24,* 303–325.

Spearman, C. (1904). "General intelligence" objectively determined and measured. *American Journal of Psychology, 15,* 201–293.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103,* 670–680.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review, 64,* 153–181.

Thurstone, L. L. (1927). Psychophysical analysis. *American Journal of Psychology, 38,* 368–389.

Wiig, E. H., & Semel, E. M. (1976). *Language disabilities of children and adolescents.* Columbus OH: Merrill.

Zimmerman, W. S. (1946). A graphic of orthogonal rotation of axes. *Psychometrika, 11,* 51–56.

Zwicky, F. (1957). *Morphological analysis.* Berlin: Singer.

Author's Address

Send requests for further information to J. P. Guilford, P.O. Box 1288, Beverly Hills CA 90213, U.S.A.

Reprints

Reprints of this article may be purchased *prepaid* for $2.50 for delivery in the U.S. or $3.00 (in U.S. funds drawn on a U.S. bank) elsewhere, from Applied Psychological Measurement, N658 Elliott Hall, University of Minnesota, Minneapolis MN 55455, U.S.A.