

# Educational and Psychological Measurement

<http://epm.sagepub.com/>

---

## **A Factor-Analytic Study of Spontaneous-Flexibility Measures**

Frank B. May and Alan W. Metcalf

*Educational and Psychological Measurement* 1965 25: 1039

DOI: 10.1177/001316446502500409

The online version of this article can be found at:

<http://epm.sagepub.com/content/25/4/1039>

---

Published by:



<http://www.sagepublications.com>

**Additional services and information for *Educational and Psychological Measurement* can be found at:**

**Email Alerts:** <http://epm.sagepub.com/cgi/alerts>

**Subscriptions:** <http://epm.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

**Citations:** <http://epm.sagepub.com/content/25/4/1039.refs.html>

## A FACTOR-ANALYTIC STUDY OF SPONTANEOUS-FLEXIBILITY MEASURES<sup>1</sup>

FRANK B. MAY AND ALAN W. METCALF

Washington State University

THE purposes of this study were (1) to determine the extent to which spontaneous flexibility is independent of certain changes in test instructions and scoring procedures, and (2) to compare the effectiveness of three different means of deriving scores of spontaneous flexibility.

### *Background and Rationale*

The phenomenon called "flexibility" has been investigated for many years (Cattell and Tiner, 1949; Guetzkow, 1951; Guilford, Frick, Christensen, and Merrifield, 1957; Kleemeier and Dudek, 1950; Luchins, 1942; Oliver and Ferguson, 1951). More recently, however, flexibility has become a popular variable in creative thinking studies. Both Torrance (1960, 1962) and Guilford (1952, 1956, 1961), for example, have utilized flexibility measures to a great extent in determining the creative potential of adults and children.

So far, however, Torrance and Guilford have differed considerably on how to measure flexibility. The measurements of the two investigators differ in important respects such as scoring procedures, mental sets induced by test instructions, and examination tasks. Furthermore, Guilford's studies indicate at least two types of flexibility—spontaneous and adaptive—whereas Torrance makes no attempt to separate the two types. The effect of fluency on Guilford's measurements of flexibility has been made evident through factor

---

<sup>1</sup> This study was part of the Spokane Dropout Study financed by the State Department of Public Instruction in Olympia, Washington, and directed by Lloyd B. Urdal, Department of Education, Washington State University.

analysis. The effect of fluency on Torrance's measurements of flexibility has not been made evident. Moreover, the effect of different test instructions on measurements of flexibility has not been systematically studied, although Torrance (1962-b) has examined such effects on measurements of originality. In another study of originality, Christensen, Guilford, and Wilson (1957) found that tests on which the subjects were instructed to be original and tests on which the subjects were *not* instructed to be original both had high loadings on the same factor. It would seem equally important to ascertain whether flexibility is similarly independent of the mental set induced by test instructions.

In addition to further psychometric delineation of flexibility, scoring procedures which are both conservative of time and reasonably objective need to be established. Torrance and Guilford have developed at least three different flexibility scores for tests of similar content. With Guilford's "Brick Uses" test, a score for spontaneous flexibility is derived by counting the different *categories* of uses employed by the subject—categories such as building uses, uses as a weight, and uses as a pounding instrument. With Torrance's "Tin Can Uses" and "Book Uses" a multi-category scheme is also used (although the measurement is called "flexibility" rather than "spontaneous flexibility"). A multi-category scheme, however, can be very time consuming and quite subjective.

In addition to the category scheme for scoring "Brick Uses," Wilson and other associates of Guilford have derived another score for spontaneous flexibility by counting the number of unusual or *alternate uses* which the subject thinks of for an object. Each item in their Alternate Uses Test (1960) presents the name of a common object, along with a statement of its most common use. The subject is asked to think of six other uses which are less common. The test consists of nine items in sets of three, with four minutes given for each set. A sample item is presented to the subject. The score for this test is based on the number of acceptable responses; "vague," "impossible," and "overworked" responses are considered unacceptable. This test eliminates some of the shortcomings of tests which must be scored by the category scheme. However, other shortcomings seem evident: (1) Limiting a subject to six responses per item may stifle his flexibility "just when he gets going." (2) Informing the subject of the common use of the object and of the necessity to

consider less common uses seems, logically, to provide too much of a clue to the subject; to measure *spontaneous* flexibility, it appears illogical to *tell* the subject to be flexible. (3) To measure *spontaneous* flexibility it seems inappropriate to show the subject *how* to be flexible by providing him with a sample item. (4) The scoring protocol is somewhat lacking in specificity.

In addition to the category scheme for scoring "Tin Can Uses" and "Book Uses," Torrance has derived another score for flexibility by counting the number of different *principles* employed by a subject in thinking of improvements for an object. Torrance has delineated 20 principles, including maximization, minification, and addition, for this scoring procedure. Essentially, this is another category scheme. However, no matter what object the subject is asked to "improve," the same twenty principles can be used to score his responses for flexibility; a new set of categories does not have to be established for each object. This scoring procedure shares the deficiency of the category scheme in that scorers frequently have difficulty classifying responses. (A more severe weakness will be discussed later.)

#### *Modified Testing and Scoring Procedures*

In an attempt to overcome some of the limitations in the testing and scoring procedures which have been discussed, modified procedures for testing and scoring spontaneous flexibility were developed by the writers. Our purpose was to develop procedures which would meet the following criteria:

1. Lack of clues to the subjects that spontaneous flexibility is called for.
2. Relatively fast and objective scoring.
3. Relatively low correlation with fluency.

Following is a test item and scoring procedure designed to meet those criteria:

#### Instructions:

List as many uses as you can think of for one or more pencils.  
You will have only three minutes.

#### Scoring Protocol (unconventional uses):

Score only one point for *marking* with the lead (writing, shading, drawing, etc.); only one point for *erasing* (no matter

what is erased); and an additional point for each other use, other than marking or erasing. *However,*

- (a) Score no point for an impossible use, but avoid a strict interpretation of "impossible." (An example of an "impossible" use would be "a cheap tool for cutting soft diamonds.")
- (b) Score no more than two points for an action verb repeated within a single item (e.g., "poking people," "poking cats.")
- (c) Score no additional point for any response which *can* be placed in a "conventional use" category, (e.g., "in arithmetic class" can be placed in the category of "marking.")
- (d) Score no point for any response which is not an actual *use* (e.g., a preparation for use, such as "sharpening;" an action which is only related to the object, such as "something to forget;" dispensation after use, such as "putting in your pocket.")

### *Method*

#### *Subjects*

The subjects of the present study consisted of 332 eighth-grade students in a Spokane, Washington high school. The students were heterogeneous with respect to scholastic abilities, sex, and socio-economic status.

#### *Tests*

The following group tests (which included only verbal stimuli) were administered to the students:

#### *Creative Thinking Battery*

1. Table Fork Improvement. This test was adapted from Torrance's "Fire Truck Improvement" (1960) and scored for fluency (total number of responses), and also for spontaneous flexibility (Torrance's principles). A mental set of fluency was encouraged by warning the subjects that they had only three minutes to list as many improvements as they could.
2. Chalkboard Improvement. (Same as number one.)
3. Pencil Uses. This test was adapted from Guilford's "Brick Uses"

(1952) and scored for fluency and also for spontaneous flexibility (unconventional uses). A mental set of fluency was encouraged by warning the subjects that they had only three minutes to list as many uses as they could.

4. Tin Can Uses (Torrance, 1960). This test was scored for fluency and also for spontaneous flexibility (Torrance's categories). A mental set of fluency was encouraged as in number three.
5. Broom Uses. (Same as number three.)
6. Book Uses. (Same as number four.)
7. Unusual Substitutes. This test, consisting of two parts, was adapted from Wilson's "Unusual Uses" (1953) and was scored for uncommonness of response. The score was derived by weighting each response from one to five depending on the frequency of its occurrence in this population. This test was used as a reference variable for "conceptual adaptive flexibility" (Guilford, 1957) and as a buffer between the fluency set and the subsequent flexibility set. The test instructions follow:

In the following test list five things which might be substituted for each object below. In other words, if you did not have the following objects available, what might you use instead? Try to think of the most *unusual* substitutes that you can. However, they must be reasonable and practical, even if they are unusual.

A mental set of flexibility was further encouraged by telling the subjects to "Take your time. Don't hurry. You will have a full seven minutes."

8. Alarm Clock Improvement. This test was adapted from Torrance's "Fire Truck Improvement" (1960) and scored for spontaneous flexibility (Torrance's principles). A mental set of flexibility was encouraged by telling the subjects to "list many *different kinds of improvements*. For example, besides improving the *sound* of an alarm clock, also list other kinds of improvements. Take your time. Don't hurry. You will have a full five minutes."
9. Rocking Chair Improvement. (Same as number eight, except no example was given.)
10. Table Knife Uses. This test was adapted from Guilford's "Brick Uses" and was scored for spontaneous flexibility (unconventional uses). A mental set of flexibility was encouraged by tell-

ing the subjects to "list many *different kinds* of uses. For example, besides listing cutting uses, also list other kinds of uses. . . . Take your time. Don't hurry. You will have a full five minutes."

11. Needle Uses. (Same as number ten, except no example was given.)

### *Differential Aptitude Tests*

Only the score on the "Verbal Reasoning Test" was used in this study. "Verbal Reasoning" is a power test consisting of 50 verbal-analogy items. It was designed to measure a combination of the "verbal ability" and "deductive reasoning" factors (Carroll, 1959) and correlates highly with intelligence scores (Frederiksen, 1959). Both the first and last components of each analogy are omitted, and the subject must choose from several alternatives the words which will complete the analogy. Because the population used in this study was heterogeneous, the Verbal Reasoning Test was included as a reference variable for verbal intelligence.

### *Treatment of Data*

To increase the range of the scores and the reliability of the inter-correlations, several pairs of scores were combined, resulting in 11 variables as shown in Table 1. A perusal of variables 4-7 will indicate how two of the scoring procedures for spontaneous flexibility were each paired with two mental sets. An obvious weakness in our design is the lack of a partner variable for variable 8. It would have been desirable to administer another pair of "Uses" tests under a mental set of flexibility, which then would have been scored for spontaneous flexibility, using the category scheme. The omission of this variable was partly a concession to expediency: a maximum of 50 minutes was allotted for the creative thinking battery; also, scoring time would have been increased considerably. The omission was also a result of theoretical considerations: Torrance had developed categories for only two tests—"Tin Can Uses" and "Book Uses;" the writers considered it appropriate to test only *his* categories and *his* principles, since Torrance seems to consider both schemes as equivalent means of scoring flexibility. The effect of this omission of a partner variable for variable 8 will be discussed in a later section.

TABLE 1  
*Means, Standard Deviations and Reliability Estimates of Eleven  
 Tests Administered to 332 Heterogeneous Eighth Graders*

Var.	Tests	Scoring	Set	Task	Mean	S. D.	$r_{11}$ *	$R$ **
Creative Thinking Tests								
1	1 + 2	Fluency	Flu.	Improvements	8.56	3.55	.80	.99
2	3 + 5	Fluency	Flu.	Uses	13.29	5.56	.74	.99
3	4 + 6	Fluency	Flu.	Uses	16.61	6.78	.62	.99
4	1 + 2	Principles	Flu.	Improvements	6.68	2.46	.62	.88
5	8 + 9	Principles	Flex.	Improvements	8.07	2.94	.77	.87
6	3 + 5	Unconv. Uses	Flu.	Uses	6.68	3.37	.50	.93
7	10 + 11	Unconv. Uses	Flex.	Uses	9.97	5.01	.77	.98
8	4 + 6	Categories	Flu.	Uses	9.72	3.74	.51	.90
9	7 a	Uncommonness	Flex.	Substitutes	12.58	3.63	.23	N
10	7 b	Uncommonness	Flex.	Substitutes	12.60	4.58	.22	N
Differential Aptitude Test								
11	V. R.	Correct Responses		Analogies	20.29	8.03	.88	N

\* Estimated for variables 1-8 by a stepped-up correlation between test scores. Estimated for variables 9 and 10 by using the communality as a low estimate. Estimate for variable 11 by using the parallel-forms coefficient reported in the test manual.

\*\* Interjudge reliability  
 N Not computed

The 11 variables were factor analyzed by means of the CDC 1604 electronic computer, using a program designed by Harris (1962). This program utilizes the principle axes algorithm, with the roots and vectors obtained from the matrix  $U^{-1}RU^{-1}$  where  $U$  is a diagonal matrix of uniqueness estimates and  $R$  is the correlation matrix.  $U^2$  is estimated by one minus the square of the multiple correlation coefficient. Factors corresponding to eigenvalues greater than 1.0 are retained. These factors are subjected to a varimax rotation.

### Results

Table 2 shows the intercorrelations of the eleven variables described in Table 1. Tables 3 and 4 show the unrotated and rotated factor matrices. Although six factors were extracted and rotated, only four will be described in Table 5. The other two factors displayed loadings which the writers considered psychologically insignificant (see Table 4), and accounted for less than 3 per cent of the total variance. The four factors described in Table 5 accounted, respectively, for 18%, 16%, 10%, and 7% of the total variance.

### Discussion

It is likely that FACTOR I is similar to the dimension referred to by Guilford (1952) as "ideational fluency." However, since variables



TABLE 2\*  
*Intercorrelations of Eleven Variables\*\**

	1	2	3	4	5	6	7	8	9	10
2	54									
3	50	76								
4	85	44	44							
5	54	44	46	52						
6	30	51	31	27	24					
7	17	32	32	18	45	28				
8	37	47	55	36	35	60	32			
9	18	18	12	17	20	18	25	15		
10	16	19	21	16	17	27	18	29	25	
11	02	09	00	10	08	27	16	22	10	13

\* Decimal points omitted

\*\* See Table 1 for description of variables

TABLE 3\*  
*Unrotated Factor Matrix for Eleven Variables\*\**

Variable	Factors					
	I	II	III	IV	V	VI
1	-.83	-.36	.00	.05	-.02	-.01
2	-.76	.31	.20	.06	-.13	-.09
3	-.74	.29	.32	-.01	.07	.03
4	-.79	-.41	-.09	.02	.02	.02
5	-.64	-.04	-.03	-.33	.05	-.04
6	-.50	.39	-.36	.18	-.07	.05
7	-.37	.28	-.14	-.41	.00	-.09
8	-.61	.38	-.22	.09	.18	.11
9	-.24	.06	-.20	-.22	-.20	-.07
10	-.27	.18	-.22	-.07	-.03	.16
11	-.13	.17	-.37	.01	-.02	-.15

\* Decimal points omitted

\*\* See Table 1 for description of variables

TABLE 4\*  
*Rotated Factor Matrix for Eleven Variables\*\**

Variable	Factors					
	I	II	III	IV	V	VI
1	-.85	.29	-.09	-.10	-.04	.02
2	-.31	.75	-.21	-.19	-.12	.00
3	-.30	.77	-.07	-.17	.09	.10
4	-.86	.19	-.12	-.11	.00	.06
5	-.47	.28	-.06	-.46	.10	.07
6	-.14	.33	-.64	-.11	-.07	.17
7	-.06	.23	-.19	-.55	.08	.06
8	-.21	.44	-.52	-.13	.20	.22
9	-.13	.03	-.17	-.35	-.16	.03
10	-.09	.12	-.27	-.18	.00	.25
11	-.01	-.04	-.41	-.16	-.01	-.07

\* Decimal points omitted

\*\* See Table 1 for description of variables

1 and 4 were not experimentally independent (the same pair of tests were scored for two different things), it is difficult to know whether Factor I involves "fluency with improvements" or "flexibility with principles." On the other hand, since Table 4 shows that several fluency measures had moderate to high correlations with this factor, and since Table 5 shows that variable 5 also had a high loading on the adaptive flexibility dimension, it seems reasonable to hypothesize that Factor I represents "fluency with improvements" rather than "flexibility with principles." Most important for this study, it appears that the "principles" scheme yields scores which are highly contaminated by the relative fluency of the subjects. Even

TABLE 5  
*Four Factors Extracted from Intercorrelations of Eleven Variables*

Variable	Scoring	Task	Mental Set	Loading
FACTOR I: Ideational Fluency (A)				
4	Principles	Improvements 1 + 2	Fluency	.86*
1	Fluency	Improvements 1 + 2	Fluency	.85*
5	Principles	Improvements 8 + 9	Flexibility	.47*
FACTOR II: Ideational Fluency (B)				
3	Fluency	Uses 4 + 6	Fluency	.77*
2	Fluency	Uses 3 + 5	Fluency	.75*
8	Categories	Uses 4 + 6	Fluency	.44
FACTOR III: Spontaneous Flexibility				
6	Unconv. Uses	Uses 3 + 5	Fluency	.64*
8	Categories	Uses 4 + 6	Fluency	.52*
11		Verbal Reasoning		.41*
FACTOR IV: Adaptive Flexibility				
7	Unconv. Uses	Uses 10 + 11	Flexibility	.55*
5	Principles	Improvements 8 + 9	Flexibility	.46
9	Uncommonness	Substitutions 7 a	Flexibility	.35*

\* Highest loading for this variable on any of the four factors

variable 5, which was independent of variable 1 in testing sequence, mental set, and scoring, had its highest loading on Factor I.

FACTOR II also appears similar to the dimension which Guilford calls "ideational fluency." It is interesting, however, that with this population two different types of ideational fluency were isolated. It appears that Factor II is a dimension which relates primarily to the task of thinking up uses for objects under the mental set of fluency. It is tempting, then, to label Factor II as "ideational fluency: uses" and to label Factor I as "ideational fluency: improvements." These labels, of course, are hunches. It is significant, for purposes of this investigation, that variable 8 (scored for categories), had a high loading on this factor, whereas variable 6 (scored

for unconventional uses), had a relatively low one (see Table 4). This suggests that "unconventional uses" as a scoring procedure was less affected by fluency than "categories" as a scoring procedure.

A comparison of FACTOR III with factors isolated by Guilford (1952, 1957, 1961) suggests that this factor is similar to the dimension which Guilford calls "spontaneous flexibility." This factor appears to be unrelated to ideational fluency and seems, rather, to be the result of the varying degree to which the subjects spontaneously shifted in their thinking from one type of use to other types of uses for an object. Both the "unconventional uses" scheme and the "categories" scheme were effective means of scoring this dimension. The "principles" scheme seems to be an ineffective means of scoring spontaneous flexibility. The high loading of the verbal reasoning score on this factor is to be expected. Several of the items in the Verbal Reasoning Test seem to require flexibility, since they contain words which have alternative meanings and require the subject to shift in his thinking in order to use the meaning which permits interpretation leading to a correct answer. These items are somewhat similar to those employed by Guilford's test called "Implied Uses" (1952), which had a high loading on the factor that Guilford labeled "spontaneous flexibility." Verbal Reasoning also may have had a high loading on Factor III because of its high correlation with general intelligence. It has been shown by Ripple and May (1962) that scores of spontaneous flexibility have a high correlation with intelligence scores when the population tested is heterogeneous.

FACTOR IV is labeled "adaptive flexibility" for two reasons: First, variable 9, a reference variable, is similar to those variables in Guilford's investigations (1952, 1956) which have been indicators of "originality." Guilford (1957) has more recently suggested that this type of "originality" would be more appropriately described as "conceptual adaptive flexibility." Second, for all three variables with a high loading on this factor, the mental set was flexibility. Since the mental set was flexibility rather than fluency, Factor IV definitely should not be considered as *spontaneous* flexibility. Both the "unconventional uses" scheme and the "principles" scheme were effective means of scoring this dimension. It seems quite possible that "categories" would have been an effective means of scoring this dimension also, if additional "uses" tests had been administered under a mental set of flexibility and scored for different categories.

*Summary and Conclusions*

The effects of different test instructions and scoring procedures on the dimension called "spontaneous flexibility" were examined by means of factor analysis. Three means of scoring spontaneous flexibility were compared to determine their relative effectiveness. Analysis was based on the test scores of a heterogeneous group of 332 eighth-grade students.

Within the limits of this study, the following conclusions are offered:

1. Spontaneous flexibility is probably *not* a factor which is independent of scoring system, mental set induced by test instructions, and examination task.
2. Spontaneous flexibility probably should be measured with tests of uses rather than improvements, under a mental set of fluency rather than flexibility.
3. Spontaneous flexibility probably should be scored either by the "unconventional uses" scheme or the "categories" scheme, but not by the "principles" scheme.
4. Adaptive flexibility can be measured either with tests of uses or tests of improvements, as long as the mental set is flexibility.
5. Adaptive flexibility can be scored by the "unconventional uses" scheme or the "principles" scheme or possibly the "categories" scheme.
6. The "unconventional uses" scheme appears to be the most economical and objective means of deriving a fluency-free score of spontaneous flexibility or adaptive flexibility.

## REFERENCES

- Carroll, J. B. in *The Fifth Mental Measurements Yearbook*, edited by Oscar K. Buros, Highland Park, New Jersey: Gryphon Press, 1959.
- Cattell, R. B. and Tiner, L. G. "The Varieties of Structural Rigidity." *Journal of Personality*, XVII (1949), 321-341.
- Christensen, P. R., Guilford, J. P., and Wilson, R. C. "Relations of Creative Responses to Working Time and Instructions." *Journal of Experimental Psychology*, LIII (1957), 82-88.
- Frederiksen, N. in *The Fifth Mental Measurements Yearbook*, edited by Oscar K. Buros, Highland Park, New Jersey: Gryphon Press, 1959.
- Guetzkow, H. "An Analysis of the Operation of Set in Problem-

- Solving Behavior." *Journal of General Psychology*, XLV (1951), 219-244.
- Guilford, J. P. "Creative Thinking Abilities of Ninth-Grade Students." A paper read at the convention of the American Educational Research Association in Chicago, February 24, 1961.
- Guilford, J. P., Frick, J. W., Christensen, P. R., and Merrifield, P. R. "A Factor-Analytic Study of Flexibility in Thinking." Reports from the Psychological Laboratory No. 18; University of Southern California, April, 1957.
- Guilford, J. P., Kettner, N. W., and Christensen, P. R. "A Factor-Analytic Study Across the Domains of Reasoning, Creativity, and Evaluation: II. Administration of Tests and Analysis of Results." Reports from the Psychological Laboratory No. 16; University of Southern California, March, 1956.
- Guilford, J. P., Wilson, R. C., and Christensen, P. R. "A Factor-Analytic Study of Creative Thinking: II. Administration of Tests and Analysis of Results." Reports from the Psychological Laboratory No. 8; University of Southern California, July, 1952.
- Harris, C. W. "Some Rao-Guttman Relationships." *Psychometrika*, XXVII (1962), 247-263.
- Kleemeier, R. W. and Dudek, F. J. "A Factorial Investigation of Flexibility." EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT, X (1950), 107-118.
- Luchins, A. S. "Mechanization in Problem Solving, The Effect of Einstellung." *Psychological Monographs*, LIV (1942), No. 6 (Whole No. 238).
- Oliver, J. A. and Ferguson, G. A. "A Factorial Study of Tests of Rigidity." *Canadian Journal of Psychology*, V (1951), 49-59.
- Ripple, R. E. and May, F. B. "Caution in Comparing Creativity and IQ." *Psychological Reports*, X (1962), 229-230.
- Torrance, E. P. "Creative Thinking of Children." *Journal of Teacher Education*, XIII (1962), 448-460.
- Torrance, E. P. "Testing and Creative Talent." *Educational Leadership*, XX (1962), 7-10+.
- Torrance, E. P., Yamamoto, K., Schenetzki, D., Palamutlu, N., and Luther, B. *Assessing The Creative Thinking Abilities of Children*. Bureau of Educational Research, College of Education, University of Minnesota, 1960.
- Wilson, R. C., Christensen, P. R., Merrifield, P. R., and Guilford, J. P. "Alternate Uses: Form A: Manual of Administration, Scoring and Interpretation (Preliminary Edition)." Beverly Hills, California: Sheridan Supply Company, 1960.
- Wilson, R. C., Guilford, J. P., and Christensen, P. R., "The Measurement of Individual Differences in Originality." *Psychological Bulletin*, (1953), 362-370.