DOCUMENT RESUME

ED 245 446 EC 162 724

AUTHOR Zarnegar, Zohreh; Hocevar, Dennis
TITLE Components of Original Thinking in Gifted Children.
PUB DATE Apr 84
NOTE 13p.; Paper presented at the Annual Conference of the American Educational Research Association (68th, New Orleans, LA, April 23-27, 1984).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cognitive Processes; *Creativity; *Gifted; Intermediate Grades

ABSTRACT

The dimensionality of original thinking in gifted children was investigated using measures designed to control for the ideational fluency/originality confounding and for task ambiguity in 315 gifted fourth to sixth graders. Ss were instructed to give only one response on each originality task rather than to give as many as possible and were encouraged, as part of the instructions to the task, to be original. Results indicated that original thinking was conceptually distinct from intelligence. However, evidence for a general originality factor was mixed. Alternative procedures for controlling the fluency/originality confounding are suggested and theoretical and methodological implications are discussed. (Author/CL)

Components of Original Thinking in Gifted Children

Zohreh Zarnegar
University of Southern California

Dennis Hocevar
University of Southern California

ABSTRACT

The dimensionality of original thinking is investigated using measures designed to control for the ideational fluency/originality confounding and for task ambiguity. Subjects were 315 gifted 4th-6th graders. Results indicated that original thinking was conceptually distinct from intelligence. However, evidence for a general originality factor was mixed. Theoretical and methodological implications are discussed.

Paper presented at the 1984 conference of the American Educational Research Association.

Despite the multitude of studies in the area of original thinking, little attention has been given to the psychometri- properties of originality measures. This is surprising in ligh· of the fact that many originality measures appear to require potentially different mental skills. For example, some measures require subjects to generate unusual uses for objects (Torrance, 1974; Guilford, 1962; Wallach & Kogan, 1965), some require subjects to speculate on how common objects might be improved (Torrance, 1974), and some require subjects to generate consequences to hypothetical situations (Torrance, 1974; Guilford, 1962). An important issue is the degree to which performance on these tasks intercorrelate. If substantial correlations between the tasks exist, then there is justification for an originality construct. However, if correlations between the various tasks are low, then the use of the originality construct is less justified. In terms of Campbell and Fiske's (1959) classic paradigm for establishing construct validity, different measures of the same construct should have convergent validity.

Since multiple measures of originality have been included in dozens of studies, it would seem that there is presently ample evidence as to whether or not originality measures have convergent validity. Unfortunately, this is not the case. Two major problems prevent researchers from making valid generalizations about the originality construct. First, there is the criticism that contemporary measures of originality are contaminated by a strong fluency component. (Fluency is defined as the number of responses given on an originality task.) Both Hocevar and Michael (1979) and Clark and Mirels (1970) have made this criticism and suggest that fluency and originality are confounded due to the additive nature of the scoring formula used in originality tests. Specifically, examinees who give more responses get higher originality scores as a result of their fluency, not their originality.

Evidence for the confounding of fluency and originality stems from a variety of analytic techniques:

a. Fluency/originality correlations are extremely high. In a review of 89 reported correlations between fluency and originality, Hocevar (1979c) found that 82% exceeded .50 and that the average fluency/originality correlation was .69. Furthermore, there is reason to believe that published studies actually underestimate the fluency/originality correlation (see discussion below).

b. Factor analytic studies do not yield separate fluency and originality factors (Child, 1970; Kazelskis, 1972; Plass, Michael & Michael, 1974; McKinney & Forman, 1977; Yamamoto & Frengel, 1966; Khattab et al., 1981).

c. Multitrait-multimethod analyses do not support the discriminant validity of the fluency/originality distinction

(Harvey, Hoffmeister, Coates, & White, 1970; Hocevar, 1979c).

d. The reliability of originality tests drops to zero or near zero when the effects of fluency are removed, thus suggesting that originality measures have little or no unique variance (Hocevar, 1979ab).

The above studies provide considerable support for the viewpoint that fluency and originality are not empirically distinct. Despite the growing consensus on this point there is little agreement on what should be done to alleviate this problem.

A second potential problem with contemporary measures of originality is that examinees are not encouraged to use the ability that the tests purport to measure. More specifically, examinees are not instructed to be original when they are providing responses, or they are given ambiguous instructions. Thus individual differences may be a matter of task perception, rather than originality. In support of this viewpoint, Harrington (1975) administered the Alternate Uses test to 50 male college students who were explicitly told to provide creative (i.e., original) uses for objects. A group of 55 control subjects were given the same test under the traditional testing mode. That is, they were not told to provide creative responses. Results conclusively indicated that the scores obtained using qualitatively oriented instructions (i.e., be creative) were better predictors of a self-report criterion measure of creativity. Despite the importance of Harrington's findings, the modifications suggested by his research have not been adopted.

Collectively, the above two measurement problems make the interpretation of prior correlations among originality measures ambiguous. In addition, prior research on the originality/intelligence issue is similarly difficult to interpret. While most researchers presently consider originality as conceptually distinct from intelligence (e.g., Torrance, 1974; Wallach & Kogan, 1965), it is possible that this conclusion may have been inaccurately brought about by the fluency/originality confounding or the ambiguous nature of the instructions on contemporary originality tasks.

The major purpose of the present study is to introduce a new way of measuring originality and to investigate the dimensionality of original thinking using measures which are not subject to the shortcomings discussed above. Two innovations will be considered. First, examinees are instructed to give only one response on each originality task rather than to give as many as possible. This modification avoids the ideational fluency/originality confounding discussed above. Second, examinees are specifically encouraged to be original as part of the instructions to the task.

Methodology

Subjects. Subjects were 315 gifted and highly gifted and talented students 4th, 5th, and 6th grade elementary gifted school programs. Of the 315 subjects, 184 were male, 131 female. 71 were 4th graders, 99 5th graders, and 145 6th graders. Giftedness was determined using the Stanford-Binet test.

Measures. The tests of originality used in this study incorporated the revisions of traditional measures of original thinking discussed above. These measures included three tests, each with 10 items. In the first test, Unusual Uses, subjects were instructed to generate an unusual use for each of 10 ordinary objects. In the second test, Improvements, subjects generated an improvement for each of 10 common objects. In the third test, Consequences, subjects generated a nonobvious consequence to each of 10 hypothetical situations. No time limit was put on the test, but the majority of the students needed approximately 30 minutes to complete the test.

Scoring. The responses to all items were subjectively scored on a four level originality scale, and total scores were computed by adding the ten items. Thus the range of possible scores was 10-30. An original answer was defined a one that was both clever and unusual. A codebook of original responses was used by the scorer. A high degree of interjudge reliability (greater than .90) in several related studies (Bachelor et al. 1980; Urman et al., 1982) provide strong evidence for the efficacy of using qualitative criteria in assessing originality.

Results

Table 1 gives the means and standard deviations for the three originality measures and for the Stanford-Binet IQ test. An initial analysis was conducted to examine whether performance on the three originality measures was related to grade level. Only the Consequences test correlated significantly with grade level, $r(313)=.17$, $p < .05$.

The intercorrelations among 30 items from the originality measures (10 each for Uses, Improvements, and Consequences) and the SB IQ measure were then factor analyzed using the LISREL confirmatory factor analysis procedure (Joreskog, 1969; Joreskog & Sorbom, 1981). Unlike traditional exploratory factor analytic techniques, confirmatory factor analysis allows for the testing of a specific factor analytic model. The model chosen for the present study is shown in Table 2. Each item for the Uses test is restricted to load on Factor one (Uses); each item from the Improvements test is restricted to load on Factor 2 (Improvements); and each item from the Consequences test is restricted to load on Factor 3 (Consequences). The Stanford-Binet IQ test defines the fourth factor - intelligence. As shown in Table 2, the CFA loadings generally support the hypothesized model in that each of the loadings is positive and significant.

Further support for the hypothesized model is obtained from the item/factor correlations (not shown). In all cases items correlated highest with the factor that they purported to measure.

Table 3 shows the factor reliabilities, and factor intercorrelations. As one would expect from the CFA loadings each of the originality measures is reliable. Factor intercorrelations are also estimated by the confirmatory factor analysis. These are given above the diagonal in Table 3. Thes correlations are corrected for measurement error. For comparison purposes, the uncorrected correlations (i.e., Pearson PM correlations) are also given in Table 3 (below the diagonal). Regardless of which set of correlations one considers, the correlations provide strong evidence for three conceptually distinct originality factors which do not correlate with intelligence.

Discussion

An initial purpose of the present investigation was to introduce a new procedure for testing originality. The procedure differs from prior procedures in that subjects are specifically instructed to be original and subjects are instructed to give a single "best" response. Results indicated that the scores derived using these innovative procedures are reliable and have a well-defined factor structure. While this finding might appear to be somewhat mundane, it is an important and essential first step in the development of alternative measures of originality.

Of greater theoretical interest are the correlations between the originality dimensions. In the introduction, it was suggested that in order to justify the "originality" construct, alternative indicators of the construct should correlate. In this study, one of the indicators (Consequences) clearly was independent of the another indicator (Improvements) and only moderately correlated with the third indicator (Uses). Given this, one can conclude that the ability to generate remote consequences has little in common with the the ability to generate uses for objects and the ability to generate improvements for objects. When one considers that the Uses and Improvements tests both tap the ability to focus on high quality answers while the Consequenses test taps the ability to generate remote associations without qualitative reflection, the lack of correlation is intuitively reasonable. The practical implication of this finding is that scores on the Consequences test should not be combined with scores on the other two measures.

The moderate correlation between the Uses and Improvements test (r=.487) does suggest that these two measures have some variance in common, and perhaps one can sum these two measures into a single originality indicator. Nevertheless, it is important to point out that much more research is needed before the scientific usefulness of such ar originality composite measure is justified. A useful first step would be to develop several more indicators of originality that would be expected to correlate with the

present indicators (i.e., Uses and Improvements). Then these indicators would need to be administered along with established indicators of closely related abilities. Two likely candidates for the latter category are ideational fluency and verbal intelligence. In the event that a confirmatory factor analysis of such a battery yields a distinct originality factor, researchers would be justified in giving such a battery of originality tests further attention.

A reasonable question is whether the results of the present study are an artifact of the selective sample. The mean intelligence test score of the sample is extremely high (mean IQ = 146 vs 100 for a national sample), and there is reduced variation on the IQ measure (SD=11.2 vs 16.0 in a national sample). Perhaps this restriction of range can explain the low correlations between intelligence and the originality measures. Nevertheless, it is important to point out that there was considerable variation on the intelligence measure. Furthermore, the selective sample probably had a smaller influence on the correlations between the originality measures because the sample was not selected using performance on originality tasks as a criteria.

The above findings and caveats notwithstanding, a major purpose of the present investigation was to introduce a new technique for controlling the fluency/originality confounding. In light of this purpose, it is important to discuss some alternative procedures for controlling for the fluency/originality confounding. One such procedure was originally introduced by Wallach & Kogan (1965), and has been used by several contemporary researchers (e.g. Milgram, 1983; Richards, 1980; Hattie, 1980). In the Wallach and Kogan procedure originality scores are computed by counting the number of original responses, and ideational fluency scores are computed by subtracting the number of original responses from the total number of responses. While this procedure does attenuate the fluency/originality correlation, its application does not warrant uncritical acceptance. First, the attenuation of the fluency/originality correlation in such studies is something of an illusion because a less reliable measure of fluency (because it is based on a difference score) is analyzed. The situation is further complicated by the fact that the fluency/originality correlation is biased downward because of the natural tendency for subtrahend (i.e, originality) to be negatively correlated with a difference score (i.e., fluency). Finally, even if one ignores these potential biases, not counting original responses in a fluency count is difficult to rationalize, conceptually.

A second approach to controlling for ideational fluency has been introduced by Clark & Mirels (1970). Under the Clark/Mirels system originality scores are computed by scoring originality from the first three responses, thereby eliminating the advantage normally given to subject's who have given a large number of responses. Conceptually, there are problems with this approach in that it seems somewhat arbitrary to penalize respondants who put

their best responses at the end of their list. Indeed, there is even evidence that earlier responses on originality tests are generally less original. Another criticism of the Clark/Mirels formulation was raised by Speedie, Asher, & Treffinger (1971) who point out that the "positive" effects of Clark and Mirels are mostly due to lowered reliabilities, thus making the correction more illusory than real.

While there are problems with both the Wallach/Kogan and the Clark/Mirels procedures for controlling for fluency and originality, there are also advantages to their procedures. First, they allow fluency and originality to be measured simultaneously. Furthermore, their procedures do not require subjects to evaluate their own responses (i.e., choose a "best" answer). Whether one or both of these procedure will prove to be superior (or inferior) to the procedures recommended in the present investigation can only be decided after much further research. In the meantime, the potential merit and problems with each procedure deserve to be carefully considered by researchers who want to investigate the fluency/originality construct(s).

References

Bachelor, P., Hocevar, D., Bachelor, B., & Hocevar, S.P. A new technique for measuring originality in 6-9 year old children. Paper presented at NCME, 1981.

Child, D. The essentials of factor analysis. New York: Holt, 1970.

Clark, P. & Mirels, H. Fluency as a pervasive element in the measurement of creativity. Journal of Educational Measurement, 1970, 7, 83-86.

Cronbach, L.J. Coefficient alpha and the internal structure of tests. Psychometrika, 1951, 16, 297-334.

Harvey, O.J., Hoffmeister, J.K., Coates, D., & White, B.J. A partial evaluation of Torrance's tests of creativity. American Educational Research Journal, 1970, 7, 359-372.

Hattie, J. Should creativity tests be administered under testlike conditions? An empirical study of three alternative conditions. Journal of Educational Psychology, 72, 87-98.

Hocevar, D. The unidimensional nature of creative thinking in fifth grade children. Child Study Journal, 1979, 32, 273-277.

Hocevar, D. Ideational fluency as a confounding factor in the measurement of originality. Journal of Educational Psychology, 1979, 71, 191-196

Hocevar, D. Multitrait-multimethod analysis of tests of divergent thinking. Paper presented at the Conference of the Western Psychological Association, 1979.

Hocevar, D. & Michael, W. Effects of scoring formual on the discriminant validity of tests of divergent thinking. Educational and Psychological Measurement, 1979, 39, 917-21.

Kazelskis, R. The convergent, divergent, and factorial validity of the Torrance Figural Test of Creativity. Southern Journal of Educational Research, 1972, 6, 123-129.

Khattab, A., Clark, F., Hocevar, D. & Benson, J. A confirmatory factor analysis of the TTCT under four testing conditions. Paper presented at the Conference of the American Educational Research Assocation, 1981.

McKinney, J. & Forman, S. Factor structure of the Wallach-Kogan tests of creativity and measures of intelligence and achievement. Psychology in the Schools, 1977, 14, 41-44.

Milgram, R. Validation of ideational fluency measures of original thinking in children. Journal of Educational Psychology, 1983, 75, 619-624.

Milgram, R., Milgram, N., Rosenbloom, G., & Rabkin, L. Quantity and quality of creative thinking in children and adolescents. Child Development, 1978, 42, 385-388.

Nunnally, J. Psychometric Theory. New York, McGraw-Hill, 1978.

Plass, H., Michael, J., & Michael, W. The factorial validity of the Torrance Tests of Creative Thinking for a sample of 111 sixth-grade children. Educational and Psychological Measurement, 1974, 34, 413-414.

Richards, R. A comparison of selected Guilford and Wallach Kogan Thinking creative tests in conjunction with measures of intelligence. Journal of Creative Behavior, 1979, 13, #3, 151-154.

Speedie, S., Asher, J. & Treffinger, D. Comment on "Fluency as a pervasive element in the measurement of creativity". Journal of Educational Measurement, 1971, 8, 125-26.

Torrance, E.P. Torrance Tests of Creative Thinking: Norms-technical manual. Lexington, MA: Ginn & Company, 1974.

Urman,H., Adler, P., & Hocevar, D. Effects of variations and instructions and scoring criteria on the measurement of original thinking. Paper presented at the 1982 Conference of the American Educational Research Association, 1982.

Wallach, M.A. & Kogan, N. Modes of thinking in young children: A study of the creativity-intelligence distinction. New York: Holt, Rinehart & Winston, 1965.

Yamamoto, K. & Frengel, B. An exploratory component analysis of the Minnesota Tests of Creative Thinking. California Journal of Educational Research, 1966, 17, 220-229.

Table 1

Means and standard deviations

Test	Mean	SD
Uses	19.96	3.89
Improvements	17.88	5.13
Consequences	13.85	3.63
SB IQ	146.80	11.19

Table 2

Factor loadings

	1	2	3	4[a]
U1	.379	.000	.000	.000
U2	.550	.000	.000	.000
U3	.421	.000	.000	.000
U4	.488	.000	.000	.000
U5	.507	.000	.000	.000
U6	.566	.000	.000	.000
U7	.427	.000	.000	.000
U8	.339	.000	.000	.000
U9	.535	.000	.000	.000
U10	.489	.000	.000	.000
I1	.000	.613	.000	.000
I2	.000	.697	.000	.000
I3	.000	.679	.000	.000
I4	.000	.693	.000	.000
I5	.000	.594	.000	.000
I6	.000	.455	.000	.000
I7	.000	.573	.000	.000
I8	.000	.567	.000	.000
I9	.000	.635	.000	.000
I10	.000	.501	.000	.000
C1	.000	.000	.449	.000
C2	.000	.000	.479	.000
C3	.000	.000	.586	.000
C4	.000	.000	.671	.000
C5	.000	.000	.468	.000
C6	.000	.000	.455	.000
C7	.000	.000	.543	.000
C8	.000	.000	.539	.000
C9	.000	.000	.750	.000
C10	.000	.000	.662	.000
SBIQ	.000	.000	.000	.900

[a] The loading of the SB IQ test was set to .900 in order to identify the CFA model. This loading is the square root of the estimated reliability (.810). See the footnote to Table 3 for further information on this estimate.

Table 3

Factor Reliabilities[a] and Factor Intercorrelations

	1	2	3	4
1. Uses	.740	.487	.306	-.100
2. Improvements	.379	.850	.084	-.086
3. Consequences	.229	.091	.820	.078
4. SB IQ	-.056	-.079	.044	.810

Note. Factor reliabilities (internal consistencies) are shown in the diagonal. The factor correlations that are given above the diagonal are corrected for attenuation due to measurement error by CFA procedures. Factor intercorrelations below the diagonal are uncorrected Pearson product moment correlations.

[a] Internal consistency reliabilities for Uses, Improvements, and Consequences are defined in the traditional fashion (i.e. Cronbach, 1952); however, the estimate of error variation for each composite is based on the CFA findings. The reliability for the SB was based on the published reliability in a national cross-sample (.91) after it has been corrected for the restricted range of IQ scores in this study (see Nunnally, 1978, p. 241).