DOCUMENT RESUME

ED 312 293 TM 014 058

AUTHOR	Green, Kathy E.
TITLE	Analysis of Cognitive Style Measures.
INSTITUTION	Johnson O'Connor Research Foundation, Chicago, IL. Human Engineering Lab.
REPORT NO	TR-1988-2
PUB DATE	Apr 89
NOTE	47p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Age Differences; *Aptitude Tests; Cognitive Measurement; *Cognitive Style; Comparative Testing; *Psychometrics; Sex Differences; Standardized Tests; *Test Reliability
IDENTIFIERS	Analytic Ability; Matching Familiar Figures Test (Kagan); Reasoning Tests; Rod and Frame Test; Stroop Color Word Test; Test Batteries; Visual Verbal Synchrony

ABSTRACT

The psychometric utility of six experimental cognitive style (CS) measures was analyzed. Examinees were 1,135 clients of the Johnson O'Connor Research Foundation who, during 1985, completed at least one of the six CS tests. Information is provided on measure reliability; relationships among CS measures; relationships with standard battery aptitude measures; and relationships to sex, age, education, and laterality. Measurement precision appears adequate for the Rod-and-Frame Test and the Matching Familiar Figures Test style score. For the Stroop Color and Word Test, the reliability was a little low but could be raised by lengthening the test. These three measures appear to be independent of each other and of the standard battery measures and may be useful additions to the battery if further research indicates that they predict vocational/avocational performance. The Verbalizer/Visualizer Questionnaire requires further investigation and development prior to use. The measure of global-analytic style did not have adequate reliability, and the measure of global-analytic ability showed overlap with current battery measures, but might prove useful as a measure of general reasoning ability. Item statistics, the calculation routine for over- and under-correction on the Rod and Frame Test, and descriptive statistics for the standard battery measures are appended. (TJH)

ANALYSIS OF COGNITIVE STYLE MEASURES

Kathy E. Green

JOHNSON O'CONNOR RESEARCH FOUNDATION
HUMAN ENGINEERING LABORATORY

Technical Report 1988-2
April 1989

Analysis of Cognitive Style Measures

Kathy E. Green

ABSTRACT

This paper reports results of the analysis of six experimental cognitive style measures. A preliminary report was prepared by Michael Windle (Statistical Bulletin 1986-6) and should be consulted for background information, as should Technical Report 1985-1, by Kathy Green, which presented a review of the cognitive style literature. This report provides information regarding measure reliability; relationships among cognitive style measures; relationships with standard battery aptitude measures; and relationships to sex, age, education, and laterality.

The primary purpose of this study was to determine the utility of six experimental cognitive style measures. Measurement precision appears adequate for the Rod-and-Frame Test and the Matching Familiar Figures Test style score. For the Stroop Color and Word Test, the reliability was a little low but could be raised by lengthening the test. These three measures appear to be independent of each other and of the standard battery measures and may be useful additions to the battery if further research indicates them to be predictive of vocational/ avocational performance. The Verbalizer-Visualizer Questionnaire requires further investigation and development prior to use. The measure of global-analytic style did not have adequate reliability, and the measure of global-analytic ability showed overlap with current battery measures but might prove useful as a measure of general reasoning ability.

CONTENTS

LIST OF TABLES

ACKNOWLEDGMENTS

The author would like to acknowledge the following for their valued comments and assistance in the inception, conduct, and presentation of this study: Dr. Richard Smith, American Dental Association; Dr. David Schroeder, Johnson O'Connor Research Foundation; and the test administrators of the Chicago, Houston, and San Francisco offices of the Johnson O'Connor Research Foundation.

BACKGROUND AND PURPOSE OF THE STUDY

Cognitive style is considered to be a person's typical mode of perceiving, thinking, remembering, and problem-solving. It is a relatively stable response tendency reflecting stable individual differences in information processing. Cognitive style is thought to extend to a variety of tasks and to be relatively stable in adults. There are numerous views of cognitive style and multiple definitions of the overriding construct, as well as of particular styles. The identified dimensions of cognitive style are heterogeneous and lack a unifying theoretical framework. Different cognitive styles vary in specificity. Some cognitive styles may equally well be considered aptitudes or skills. Also, measures of cognitive style for which the associated concepts are conceptually clear may be only indirectly related to these concepts and therefore lack validity. Shipman and Shipman (1985) list some of the more serious problems in measuring cognitive style, including: (a) use of an achievement or ability format (rather than style), (b) confounding with other styles or abilities, (c) lack of reliability, and (d) poor cross-task generality. Needed research includes differentiation between cognitive styles and abilities, including theoretical development of the causal effect of style (if any) in the development of ability, assessment of relationships among cognitive styles, effects of providing instruction in different cognitive styles, and generality of styles across contexts.

Five dimensions of cognitive style that appeared to be conceptually distinct were identified (Technical Report 1985-1) as part of the efforts of the Johnson O'Connor Research Foundation (JOCRF) to identify new aptitudes not already part of the standard battery. One additional potential dimension was included (measured by Object Assembly) that is thought to be related to two of the selected dimensions. This report addresses the psychometric value of these experimental measures, the relationships among them, and their relationship to the aptitudes assessed by the current battery.

COGNITIVE STYLE DIMENSIONS AND MEASURES

Field independence-dependence

Field independence-dependence (FID) is the most researched cognitive style dimension. FID was originally defined as the tendency to identify correctly the true vertical in a misinformative context. An extended definition of FID is the tendency or ability to make accurate judgments independent of the context. Field independence requires reliance on an internal frame of reference. The tendency to be context-sensitive is called field dependence. FID is considered by some to be a visual-spatial ability rather than a style.

The Rod-and-Frame Test and the Embedded Figures Test are the most commonly used measures of the construct field independence-dependence. The Rod-and-Frame Test was used as an experimental measure in this study. The portable Rod-and-Frame Test (RFT; Worksample 673AA) was developed by Herman Witkin and associates at the Educational Testing Service (Oltman, 1968). In the RFT the examinee is presented with a line (the "rod") surrounded by a square frame that is tilted at an angle. The examinee is asked to adjust the line until it is vertical. The examinee must attempt to ignore the frame in order to perform the task. The RFT was used in exploratory research at the Foundation in the 1970s (see Minutes, 1975) and showed a moderate correlation with the Wiggly Block test (r = .40) and low correlations with other worksamples. Field independence-dependence has not been studied in relation to other cognitive style dimensions except for distractibility, with which it is uncorrelated (Technical Report 1985-1).

The test consists of eight trials. In the first two, the frame is shifted 28° to the left, and then for the next two trials, the frame is moved 28° to the right. The rod is moved 28° to the left for the first and third trials and then 28° to the right for the second and fourth trials. The sequence of four trials is then repeated. On each trial, the examinee is asked to align the rod to true vertical. The score for a trial is the number of degrees of deviation from true vertical. The direction of deviation is also noted. The total score for the test is the sum across the eight trials of deviation from the objective vertical.

Values are also calculated for overcorrection and undercorrection. Undercorrection is deviation from the vertical in the direction of the frame; overcorrection is deviation from the vertical away from the direction of the frame. For example, if the frame is 28° left, a left deviation from true vertical constitutes undercorrection, and a right deviation is overcorrection. The calculation routine used in this analysis is provided in Appendix A.

Distractibility

This dimension involves the degree to which individuals selectively attend to relevant stimuli and ignore irrelevant stimuli. Distractibility has been found to be unrelated to field independence-dependence and reflectivity-impulsivity. A distractibility test was first introduced in the United States by John Stroop in 1935: the Stroop Color and Word Test. In the Stroop test, after baseline measurement of word reading speed and color naming speed, the examinee is presented with a sheet on which the words red, green, and blue are printed in nonmatching ink (e.g., the word red is printed in blue ink), listed repeatedly in random sequences. The examinee's task is to name the color of ink each word is printed in, while ignoring the word itself. The test is challenging because the tendency to read printed words is so strongly trained that examinees find it difficult to inhibit it in order to name the color of the print.

The measure used is the difference between predicted (based on word reading and color naming speeds) and actual color-word scores. This difference is called the color-word residual (CWR). The following formula is used to calculate color-word residuals:

$$\text{CWR} = \text{Color-word score} - \frac{\text{Color score} \times \text{Word score}}{\text{Color score} + \text{Word score}}$$

Residual scores (CWR) are then standardized. Higher values of color-word residuals represent less susceptibility to interference.

Reflectivity-Impulsivity

Reflectivity-impulsivity is another well-researched dimension of cognitive style. It is defined as the tendency to reflect upon the accuracy of a decision or solution when presented with several alternatives. It has been suggested that individuals with an impulsive cognitive style are at a disadvantage academically (Kagan, Pearson, & Welch, 1966; Messer, 1970). Reflectivity-impulsivity has been found to be independent of distractibility (Boyden & Gilpin, 1978).

The 12-item Matching Familiar Figures Test (MFFT) is the most commonly used measure of reflectivity-impulsivity for adults. With this test, the examinee looks at the standard (a familiar figure such as a lion or a flower) and selects the standard's replicate from among eight variants.

Items are scored for both speed and accuracy, and then speed and accuracy scores are then combined into an ability score and a style score as follows: Items are scored for accuracy as 0-1 (wrong-right). Times are recorded in hundredths of minutes. Accuracy and time raw scores consist of the sum of the item scores, with the time scores reflected to represent speed. Speed and accuracy scores are standardized, and ability is then computed as the sum of speed and accuracy scores, and style as the difference between speed and accuracy. Higher scores on ability indicate greater ability. Higher scores on style indicate a reflective style, and lower scores indicate an impulsive style.

Visual-Verbal Style

A measure of visual-verbal style called the Verbalizer-Visualizer Questionnaire (VVQ) was derived by Alan Richardson at the University of Western Australia from an instrument developed by Paivio in his 1971 book Imagery and Verbal Processes. This scale measures an individual's preference for verbal versus visual thinking. Scores have been found to relate to breathing pattern and lateral eye movements (Richardson, 1977; Spoltore & Smock, 1983).

The Verbalizer-Visualizer Questionnaire consists of 15 true-false questions concerning use of or skill with verbal material and visual imagery. Items are scored as 0-1 (false-true). Items 1, 3, 4, 6, 8, 9, 10, 12, and 13 are then recoded so that, for all items, 0 represents the verbal answer and 1 represents the visual answer. The total score is the sum of the recoded items, and higher scores indicate a visual preference, while lower scores indicate a verbal preference.

Object Assembly

This test is one of the performance subtests from the Wechsler Adult Intelligence Scale-Revised. It measures a type of spatial aptitude but is also hypothesized to favor (a) individuals with a global cognitive style and (b) individuals who tend to be impulsive rather than reflective. In this test the examinee is asked to assemble each of four puzzles (a manikin, a facial profile, a hand, and an elephant) as quickly as possible. Items (puzzles) are typically scored for both speed and accuracy. Shorter times are thought to be characteristic of a global style, with longer times indicating an analytic style. Similarly, impulsive persons are hypothesized to respond more quickly than reflective persons. Accuracy scores are based on correct juxtaposition of puzzle pieces. Scoring directions were obtained from the WAIS-R manual (Wechsler, 1981), but in this analysis accuracy scores were not used since they were not always reliable and consistent across test administrators. Raw time scores were used. People who exceeded the time limit for an item were given a time score of the maximum allowed for that item (2.00 minutes for Items 1 and 2 and 3.00 minutes for Items 3 and 4). Time scores were then reflected by subtracting each score from the maximum. This was done to provide values that increased with more skilled performance.

Global-Analytic Style

In spite of considerable research in this area, the only measures extant in the literature of global-analytic problem-solving style have been content-specific, for example, problem-solving style in physics. Attempts to develop general measures have not been successful. Consequently, a general test of problem-solving style was developed in-house by drawing on published materials. The test consists of thirteen problems: six analytic problems, which can be solved most readily by systematic, step-by-step reasoning, and seven global problems, which depend for their solution on particular insights. Presumably, global problem-solvers perform better on global items, and analytic problem-solvers perform better on analytic items.

Items are scored for both speed and accuracy. For both global and analytic items, longer time scores indicate less facility with that item type. Higher accuracy scores indicate more facility with that item type. For this analysis, items were scored 0-1 (wrong-right) for accuracy, and in hundredths of minutes for time. A combined time/accuracy score ranging from 0 to 3 points was then defined for each item (see later discussion for details). As with the MFFT, two scores were calculated, an ability score and a style score. The ability score was calculated by summing the time/accuracy scores for the analytic items and standardizing, summing the time/accuracy scores for the global items and standardizing, and then summing the two scores together. The style score was calculated by taking the difference of the two scores, subtracting the global standard score from the analytic standard score. Higher style scores, then, indicate an analytic as opposed to global style. Prior to this, however, items were examined for psychometric appropriateness.

RESEARCH QUESTIONS ADDRESSED IN THIS STUDY

The overall purpose of this endeavor was to assess the psychometric value of each of the six measures used and their relationships to the standard battery measures and to each other. Additional specific questions stemming from previous research are listed below.

FID: (a) Is there a significant difference in the FID of persons in socially oriented versus analytically oriented fields? Previous research suggests socially oriented persons to be more field dependent. (b) Is FID correlated with spatial ability? Does FID as measured by the RFT form a unique factor separate from spatial ability as defined by Object Assembly, Paper Folding, and Wiggly Block? (c) Is there a sex difference in RFT scores, with males performing in a more field independent manner? and (d) Are right-handed persons and those with a strong eye preference more field independent?

Distractibility: Does the CWR from the Stroop test form a unique factor when analyzed with other cognitive style measures?

Reflectivity-impulsivity: (a) Are the norms and reliabilities on the MFFT consistent with those of previous research? (b) Is reflectivity-impulsivity as measured by the MFFT related to analytical reasoning?

Visual-verbal style: Does visual-verbal preference predict ability with verbal tests (i.e., vocabulary) and with spatial tests?

Object Assembly: Does Object Assembly score predict spatial aptitude, global problem-solving style, or impulsivity?

Global-Analytic style. Can this style be reliably measured by a general rather than a subject-specific test?

METHOD

Examinees

The examinees in this study were clients of the Johnson O'Connor Research Foundation who were tested in the Chicago, Houston, and San Francisco offices during the spring, summer, and fall of 1985. These persons came to the Foundation for testing in order to obtain information about their aptitudes useful in career and educational planning. They paid a fee for the testing. Examinee ages ranged from 14 to 65 years (mean = 27); 52% were female and 48% were male. Reported years of education ranged from 0 to 22. A total of 1,135 persons completed at least one of the six cognitive style tests. The sample was diverse with respect to both age and education.

Procedure

Testing commenced in the spring of 1985 and was completed in late fall. In July, initial data from the Global-Analytic test were analyzed, and several items were revised and/or replaced. Only data from the revised version will be presented here. Tests were given by test administrators in the San Francisco and Houston offices. In Chicago, because of a shortage of test administrators, two tests were administered by Research Department staff (VVQ and Global-Analytic). The other tests were given by test administrators. The RFT and CWR tests were usually administered in that order in the first individual appointment, after the standard tests were given. The MFFT and Object Assembly were generally given in that order in the second individual appointment. The VVQ and Global-Analytic tests were given in the experimental (Group III) session. Due to time constraints and because data from the first version were not used in the ultimate analysis, fewer persons took the Global-Analytic test. Only a subset of examinees were tested using the RFT as well, due to unavailability of the apparatus in one of the locations. Items on the Global-Analytic test were administered in different orders in Houston, Chicago, and San Francisco.

The RFT took approximately 8 minutes to administer. The CWR test and VVQ each took approximately 5 minutes. The Object Assembly test took 10 minutes, and the MFFT and Global-Analytic tests each took between 10 and 25 minutes. All tests except the VVQ were individually administered. Only the VVQ was given in a paper-and-pencil format.

Methods of Analysis

Alternative scoring systems were investigated with two of the measures (Object Assembly, Global-Analytic test). The scoring system that provided the highest internal consistency and yet was the most clearly interpretable was selected as the final scoring system. The different systems tried are mentioned in the Results section.

Analysis of all measures included examining the fit of items to a homogeneous scale, using traditional item analysis. The internal consistency reliability was calculated for each measure, as were item-total correlations. If an item appeared to misfit, omission of the item from the total scale was considered. Further analysis included a principal components analysis of the item set where appropriate. Varimax rotation was used for multi-factor solutions. If additional analysis seemed necessary, a Rasch analysis was performed using BICAL (Wright, Mead, & Bell, 1980). For several measures, some items were omitted and total scores recalculated based on the reduced item set. In other cases, items of marginal value were included, but in the future, scale revision might be considered.

Once a scale (overall) score was calculated for each measure, Pearson product-moment correlations were calculated among the cognitive style measures and between the cognitive style measures and the battery variables. Differences in cognitive style in relation to sex, age, educational level, and laterality were assessed using *t*-tests and analyses of variance. Principal components analyses were conducted to determine the factor structure of the cognitive style tests alone and as part of the standard battery. Finally, multiple regression analyses were conducted to determine whether style could be predicted from measures in the current standard battery.

The $p < .01$ level was set as the acceptable probability of Type I error, except where noted (for the college major analysis). This level was used rather than $p < .05$ due to the relatively large number of cases available. The "canned" computer package SPSS-X (SPSS Inc., 1986) was used to perform all analyses except for Rasch item analyses.

RESULTS

Measure Development

Rod-and-Frame Test. Analysis of this measure of the FID construct consisted of calculating its internal consistency reliability, performing a factor analysis, and transforming raw scores to standard scores. Examination of item-total correlations did not suggest any trial to be ineffective; the scale alpha was .93. The first trial correlated the lowest with total score, possibly due to a start-up or acclimatization effect. A principal components analysis of the item correlation matrix suggested one dominant factor accounting for 67.9% of the variance (eigenvalue of 5.43). This indicates that all items are measuring the same underlying trait. Since the items were homogeneous, a total scale score was calculated. Total scores were then reflected so that a higher score indicates a higher level of field independence and a lower score represents a higher level of field dependence. This was done by subtracting the examinee's raw score from the highest score obtained in the sample (217). Scores were then standardized by subtracting the sample mean and dividing by the sample standard deviation. Table 1 presents descriptive statistics for raw scores on the tests.

Under- and overcorrection (see earlier definitions) were summed across trials and averaged by dividing the sum by the number of trials (eight). The internal consistency reliability estimate for undercorrection was .78 and for overcorrection .67. Descriptive statistics for over- and undercorrection are also presented in Table 1.

<u>CWR</u>. Scores on this measure of distractibility were calculated using the formula presented earlier and were then standardized. Descriptive statistics are presented in Table 1 for raw scores. A reliability estimate could not be directly obtained from one test administration. The reliability given in Jensen and Rohwer (1966; .70) was taken as the estimated measure reliability.

<u>MFFT</u>. The internal consistency reliability for this measure of reflectivity-impulsivity for the time score alone was .92 and for the accuracy score alone .72. These reliability estimates are consistent with those found in previous research (Messer, 1976). The first test item had the lowest item-total correlation for both time and accuracy, indicating a start-up or acclimatization effect. Fit of the accuracy (0-1) scoring of items to a linear model was then investigated using fit statistics calculated with BICAL and via a principal components analysis. Item 1 was misfitting (total fit = 3.17). This seemed to be due to its lower item-total correlation. Item 6 was overfitting (total fit = -4.80), the overfit due to a high item-total correlation. Principal components analyses of the time and accuracy scores indicated that a single factor explained the correlations among the items. Thus, both the BICAL and principal components analyses show the items to be acceptably homogeneous and to be measuring a unitary trait.

All items were used in calculating a scale score for reflectivity-impulsivity since the item misfit was not severe and since internal consistency reliability was not lowered by inclusion of Item 1. In future use, the test administration might be altered so that test administrators time responses on one of the practice items, or an additional practice item might be included. The purpose of this would be to alleviate acclimatization effects.

Table 1

Descriptive Statistics for Cognitive Style Measures

Measure	No. items	Mean	SD	N	Skewness	Kurtosis	Range	Alpha relia.
Rod-and-Frame raw score	8	18.99	21.06	642	4.8	34.3	0,217	.93
Undercorrection/trial		1.78	2.53	642	5.3	41.2	0.27.1	.87
Overcorrection/trial		.55	.70	642	7.1	97.7	0,11.5	.69
Color-word residual raw score	-	.73	7.17	1076	2.0	.0	-40.8, +36.6	.70[a]
MFFT-time	12	11.08	5.43	1027	.6	.6	.74, 40.20	.92
-accuracy	12	7.28	2.81	1044	-.3	-.8	0,12	.72
-ability	-	.30	.89	1017	2.7	28.8	-2,+11	.57
-style	-	-.22	1.74	1017	-1.6	6.5	-14,+2	.89
Verbalizer-Visualizer	11	6.16	2.10	981	-.2	-.1	0,11	.56
Object Assembly -time	4	3.02	.97	1063	-.7	-.1	0,4.79	.48
-accuracy	4	27.55	2.93	1081			9.32	.27
Global-Analytic -analytic/item	6	1.06	.62	605	.7	.1	0,3	.59
-global/item	6	.88	.57	605	.9	.7	0,2.9	.54
-ability	-	-.01	1.70	605	.7	.1	-4,+6	.72
-style	-	.00	1.05	605	.1	.7	-4,+4	.03

Note. Higher scores represent greater field independence (RFT), less susceptibility to interference (CWR), greater ability on the MFFT, a more reflective style on the MFFT, a visual preference (VVQ), greater ability on the Global-Analytic test, and an analytic rather than global preference on the Global-Analytic test.

[a]Reliability estimate from Jensen and Rohwer (1966, p. 50).

The time scores were reflected so as to represent speed of response rather slowness. Speed correlated -.65 with accuracy. Speed and accuracy scores were then standardized and summed to form an ability scale and differenced to form a style scale, as indicated previously. The reliability of the ability score was .57; the reliability of the style score was .89. Table 1 presents descriptive statistics for raw time and accuracy scores as well as for ability and style scores. The mean time for test completion is close to that found by Heckel, Hiers, Laval, and Allen (1980) for a sample of university undergraduates.

Verbalizer-Visualizer. With all items included, the internal consistency reliability of the Verbalizer-Visualizer scale was .54. Four items (5, 8, 9, and 14) had low item-total correlations and detracted from the scale alpha. When these items were dropped, the scale alpha rose to .56, but two additional items (7 and 10) were identified that had low item-total correlations. Alpha was recalculated eliminating those two items as well, and the scale alpha rose to .58 for the remaining nine-item set. A principal components analysis of all items found four significant factors with eigenvalues greater than 1.0, accounting for 47.8% of the variance. Items loading the highest on these factors were 2, 5, 6, 11, 13, and 15 (Factor 1), 1, 3, 4, and 12 (Factor 2), 8 and 9 (Factor 3), and 7, 10, and 14 (Factor 4). The internal consistency reliability for the Factor 1 items as a set was .70 and for the Factor 2 items as a set was .67. A second principal components analysis was performed omitting the six items detracting from the scale alpha. Two significant factors were found, accounting for 48.1% of the variance. Items loading on Factors 1 and 2 were the same as found in the initial principal components analysis. These two factors were clearly interpretable as items mentioning dreams/images (Factor 1) and items mentioning words (Factor 2).

The results of these analyses suggest that more than one construct is measured by this item set. This notion is further supported by the lack of a significant correlation between the score for the visual items and the score for the verbal items, which indicates that likings for verbal and visual activities are not bipolar (inversely related), as the VVQ rationale assumes, but rather independent. This finding made it unlikely that items covering both areas could be satisfactorily fit onto a single scale. Nevertheless, further item analyses were performed in order to see if the scale functioning could be brought up to an adequate level.

First, since the results of the analyses to this point did not suggest that the items formed a homogeneous set, BICAL was used to perform Rasch analyses to identify misfitting items. When all items and all persons were included in the analysis, person separability was .55, with three items having total fit of more than |3.0|. When 86 misfitting persons were deleted, the person separability rose to .57, but item misfit was still severe. The three items with high total fit were deleted. When this was done, three additional items failed to fit the model. An analysis was then conducted eliminating all items that detracted from internal consistency reliability (Items 5, 7 through 10, and 14). Two items then overfit, and eight of the remaining nine items had between fit of more than +3.0. The person separability index dropped to .48. Similar results were found when misfitting persons were deleted.

Item-total correlations ranged from .00 to .32 for the total item set. An analysis was then performed eliminating two items with very low point biserial correlations (Items 8 and 9). Items 5 and 14 then misfit, and the analysis was rerun eliminating those items as well. Person separability was then .55, with acceptable total fit but high between fit. High between fit indicates that the items overfit--that is, the pattern formed is too close to a Guttman scale. The items function too well to fit a probabilistic model. The results of the BICAL and principal components analyses indicate that either (a) the construct measured is not well-measured or (b) more than one construct is being assessed by this item set, rather than opposite poles of the same construct. It should be noted that the "visual" items were endorsed by more persons than the "verbal" items. The differences in item "difficulties" may have contributed to the failure of the items to measure a single construct.

In any event, the 11 items with total fit values closest to zero were treated as a scale. Descriptive statistics for this measure are presented in Table 1. Revision of the measure is necessary if the scale is to be used in the future. Clarification of the relationship between the item set and the construct measured is needed.

Object Assembly. The internal consistency reliabilities of time and accuracy scores were calculated and were respectively .47 and .27. Since the distribution of time scores for Object Assembly was skewed, a square root transformation was performed for the time score. Little effect was observed on the correlation with other measures, and this transformed measure was not used further. Scores were then calculated using the time scoring described earlier, and these scores were used in all subsequent analyses. The internal consistency reliability using this scoring system was .48. Descriptive statistics for time and accuracy scores are presented in Table 1.

Global-Analytic scale. The internal consistency reliability for analytic item time scores was .49 and for global time scores .57. The correlation between analytic and global time scores was .54. The reliability for analytic accuracy scores was .81 and for global accuracy scores .43. The correlation between analytic and global accuracy scores was .83. Thus, analytic and global ability do not represent opposite poles of a continuum, as had been thought, but rather abilities that have a substantial overlap (i.e., that represent almost the same trait). This means that high analytic/low global and low analytic/high global styles do not occur very often, and are likely to be difficult to measure when they do occur.

When all of the items were treated as a single set, the reliability for time was .72 and for accuracy .63. The Horse and Rider item detracted slightly from the internal consistency of the overall scale for accuracy.

Accuracy scores (0-1 scoring) were factor analyzed, putting global and analytic items together. Two significant factors were found (eigenvalues = 5.6 and 1.0), accounting for 47.4% of the variance. Items loading substantially on Factor 1 were the Letters, Logic, Anagram, Points, Twos, Barnyard, Sayings, River, Coins, and Clock items. Items loading on Factor 2 were the Hunter and Ship items. The Horse and Rider item did not load substantially on either Factor 1 or Factor 2. Items, then, did not factor along global-analytic dimensions. This result does not support use of this scale as a measure of global-analytic style.

A scoring scheme combining time and accuracy was devised, allowing scores of 0 to 3 for each item. (Maximum allowed time per item, in hundredths of minutes, is indicated in parentheses.) Cutoff points for time scores were chosen to split examinees correctly answering the item into three roughly equal groups. The scoring rules are given in Table 2.

Internal consistency reliabilities were calculated using this scoring system and were .68 for the total scale, .50 for the global items, and .59 for the analytic items. The correlation between scaled global and analytic scores was .55. The Horse and Rider global item lowered the reliability of both the total and global scales. Recalculation without this item gave internal consistency reliability coefficients of .69 for the total scale and .54 for the global scale. The Horse and Rider item was subsequently dropped from all further analyses.

The analytic subscale, then, was formed by summing scores across the six analytic items and the global subscale by summing across the six remaining global items (eliminating the Horse and Rider item). Style and ability measures were formed by standardizing subscale scores, summing the two to create an ability measure, and subtracting global from analytic scores to form a style measure. Thus, for the style measure, high scores indicate a global style while low scores indicate an analytic style. Descriptive statistics for raw analytic and global subscale scores per item and for ability and style scores are presented in Table 1. The reliability of the ability score (a sum) was .72, not too far from acceptability (.80 or higher); however, the calculated reliability of the style score (a difference score) was near zero, .03, and so statements about Global-Analytic style are rather tenuous.

Table 2

Scoring System for Global-Analytic Problem-Solving Test

Item	Time/Accuracy	Score	Item	Time/Accuracy	Score
Letters	low - 50	3	River	low - 50	3
(150)	51 - 100	2	(90)	51 - 80	2
	101 - high	1		81 - high	1
	incorrect/no answer	0		incorrect/no answer	0
Horse, Coins,	low - 40	3	Barnyard	low - 100	3
Twos	41 - 100	2	(180)	101 - 160	2
(120,90,90)	101 - high	1		161 - high	1
	incorrect/no answer	0		incorrect/no answer	0
Logic	low - 35	3	Saying 1	low - 10	3
(90)	36 - 60	2	(60)	11 - 30	2
	61 - high	1		31 - high	1
	incorrect/no answer	0		incorrect/no answer	0
Hunter	low - 70	3	Sayings	low - 20	3
(120)	71 - 110	2	2 and 3	21 - 50	2
	111 - high	1	(60)	51 - high	1
	incorrect/no answer	0		incorrect/no answer	0
Anagram	low - 30	3	Clock	low - 80	3
(90)	31 - 70	2	(120)	81 - 120	2
	71 - high	1		121 - high	1
	incorrect/no answer	0		incorrect/no answer	0
Ship	low - 60	3	Points	low - 60	3
(90)	61 - 100	2	(90)	61 - 120	2
	101 - high	1		121 - high	1
	incorrect/no answer	0		incorrect/no answer	0

Relationships Among Measures and Factor Structure

Pearson product-moment correlation coefficients were calculated among the six style (RFT, CWR, MFFT-style, VVQ, Object Assembly, Global-Analytic style) and two ability (MFFT-ability and Global-Analytic ability) measures. Table 3 presents these correlations. The style measures were correlated at a low level and appear to be essentially independent of each other. For the style measures, the only moderate correlation was between RFT and Object Assembly (r = .36). As will be reported later, both of these tests were related to structural visualization. Object Assembly was not related to Global-Analytic style or to impulsivity, as had been hypothesized.

A higher correlation was found between ability on the Global-Analytic test and Object Assembly. Global-Analytic ability score was also significantly correlated with VVQ score (r = -.36). This correlation may be due to a joint relationship with vocabulary (see later section).

A principal components analysis with pairwise deletion and varimax rotation of components (factors) was conducted with the six cognitive style measures (excluding MFFT-ability and Global-Analytic ability). Three significant factors emerged, accounting for 60% of the variability. Object Assembly, RFT, and MFFT-style formed the first factor, with factor loadings of .73, .72, and .53, respectively. RFT, then, did not emerge as a unique factor as had been hypothesized. Global-Analytic style and CWR formed the second factor with loadings of .83 and .67, respectively. CWR did not form a unique factor either (as had been hypothesized). Visual-verbal preference formed the third factor, with a factor loading of .82. MFFT-style did not have a factor loading greater than .40 on any of these three factors when listwise deletion was used. With pairwise deletion, MFFT-style's factor loading on Factor 1 increased from .38 to .53, though it then also had a factor loading of -.34 on Factor 3. MFFT-style did not, then, clearly load on any one factor. It appears to be relatively independent of the other style variables.

The low correlations among style measures along with the results of the principal components analysis suggest that probably four different style variables are assessed: MFFT-style, VVQ, RFT, and CWR. The convergent and discriminant validities of the Global-Analytic style measure are indeterminate because of its low reliability, and Object Assembly appears to be a general IQ measure.

Relationship of Cognitive Style Measures to Battery Variables

A number of moderate to high correlations were found between cognitive style measures and battery variables. Table 4 presents all correlations significant at the .01 level, along with the multiple correlations between the battery variables and the style

Table 3

Correlations among Cognitive Style Measures

Measure	2	3	4	5	6	7	8
1. Rod-and-Frame Test (z-score)	14 639 (11)	-	13 599 (12)	-	36 612 (24)	27 334 (22)	-
2. Color-word residual (z-score)		14 995 (09)	-	14 933 (09)	-	-	--[a] 595 (15)
3. MFFT-ability			-	-	25 990 (13)	28 566 (18)	-
4. MFFT-style				-	26 990 (17)	24 566 (19)	-
5. Verbalizer-Visualizer					-	-36 592 (-23)	-
6. Object Assembly						60 591 (35)	-
7. Global-Analytic -ability							-
8. Global-Analytic -style							

Note. All correlations significant at $p < .01$ are listed. Table entries are correlations corrected for attenuation, the number of cases the correlation is based on, and the uncorrected correlation (in parentheses). Decimal points have been omitted. Higher scores represent greater field independence (RFT), less susceptibility to interference (CWR), greater ability on the MFFT, a more reflective style on the MFFT, a visual preference (VVQ), greater ability on the Global-Analytic test, and an analytic rather than global preference on the Global-Analytic test.

[a] Estimate of disattenuated correlation between Color-word residual and Global-Analytic style was unstable due to the very low reliability of global-analytic style.

Table 4

Correlations with Standard Battery Measures

Measure	GR 1	CP 2	ID 3	FO 4	IR 5	AR 6	WB 7	PF 8	PS 9	TM 10	PD 11
Rod-and-Frame Test _z_-score	-	-	-	-	-	24 531 (19)	31 552 (25)	32 552 (28)	-	15 554 (14)	17 554 (15)
Overcorrection	-	-	-	-	-	-	-	-	-	-	-
Undercorrection	-	-	-	-	-	-28 563 (-21)	-33 563 (-26)	-32 561 (-27)	-	-16 563 (-14)	-17 563 (-14)
Color-word residual _z_-score	08 925 (08)	-	-	-	13 911 (12)	11 891 (09)	13 919 (11)	16 923 (15)	-	13 925 (12)	-
MFFT-ability	22 870 (16)	-	12 865 (09)	14 867 (10)	27 858 (19)	23 834 (14)	22 865 (14)	28 867 (19)	-	15 870 (11)	-
MFFT-style	-	-	14 865 (13)	-	-	25 834 (19)	25 865 (20)	28 867 (24)	-	14 870 (13)	10 870 (08)
Verbalizer-Visualizer	-	-	-16 835 (-12)	-14 835 (-10)	-	-23 807 (-14)	-	-13 838 (-09)	-	-	-
Object Assembly	14 908 (09)	-	-	15 905 (10)	52 895 (34)	59 869 (33)	73 902 (43)	70 904 (44)	-	24 908 (16)	24 908 (15)
Global-Analytic -ability	25 551 (20)	-	17 549 (14)	-	24 544 (19)	68 526 (47)	62 549 (45)	71 551 (55)	-	25 551 (20)	32 551 (24)
Global-Analytic -style[a]	 551 (16)	-	-	-	 544 (11)	 526 (13)	-	-	-	-	-

(cont.)

Table 4 (cont.)

Measure	RM 12	MD 13	SI 14	NM 15	OB 16	FD 17	TD 18	NS 19	EV 20	MV 21	RS 22
Rod-and-Frame _z_-score	-	25 554 (21)	-	-	-	-	12 540 (11)	23 394 (21)	-	29 285 (25)	-
Overcorrection	-	-	-	-	-	-	-	-	-	-20 292 (-15)	-
Undercorrection	-	-26 560 (-21)	-	-	-15 563 (-11)	-	-11 549 (-10)	-23 395 (-20)	-	-25 292 (-21)	-
Color-word residual	-	16 924 (14)	-	12 919 (11)	10 829 (08)	-	-	-	-11 925 (-11)	-	-
MFFT-ability	21 870 (14)	30 870 (20)	-	-	23 778 (14)	21 856 (15)	-	24 654 (17)	-	-	-
MFFT-style	-	27 870 (22)	12 863 (11)	-	13 778 (10)	-	-	21 654 (18)	29 870 (27)	21 487 (18)	22 192 (18)
Verbalizer- Visualizer	-	-	-17 832 (-12)	-	-	12 822 (08)		-	-59 838 (-43)	-25 482 (-17)	-36 202 (-23)
Object Assembly	20 918 (12)	61 907 (37)	32 901 (21)	21 902 (14)	35 811 (19)	25 891 (16)	12 888 (08)	42 691 (27)	28 908 (19)	34 504 (21)	32 191 (19)
Global-Analytic -ability	31 551 (23)	50 551 (37)	37 549 (30)	37 549 (30)	25 471 (17)	-	-	67 532 (53)	49 551 (41)	65 270 (50)	-
Global-Analytic -style	-	-	549 (16)	549 (19)	-	-	-	532 (15)	-	270 (21)	-

(cont.)

Table 4 (cont.)

Measure	RC 23	RE 24	Multiple R
Rod-and-Frame Test z-score	-	-	35
Overcorrection	-	-	
Undercorrection	-	-	
Color-word residual z-score	-	-	26
MFFT-ability	-	-	31
MFFT-style	35 193 (29)	12 683 (10)	36
Verbalizer-Visualizer	-59 203 (-39)	-53 606 (-34)	49
Object Assembly	55 192 (34)	27 679 (16)	56
Global-Analytic -ability	-	48 524 (35)	70
Global-Analytic -style	-	-	23

Note. Numbers refer to battery aptitude measures as follows:

1. Graphoria percentile score
2. Color Perception raw score
3. Ideaphoria raw score
4. Foresight raw score
5. Inductive Reasoning raw score
6. Analytical Reasoning raw score
7. Wiggly Block raw score
8. Paper Folding raw score
13. Memory for Design raw score
14. Silograms raw score
15. Number Memory raw score
16. Observation raw score
17. Finger Dexterity raw score
18. Tweezer Dexterity raw score
19. Number Series raw score
20. English Vocabulary scale score

(cont.)

Table 4 (cont.)

9. Personality raw score
10. Tonal Memory raw score
11. Pitch Discrimination raw score
12. Rhythm Memory raw score
21. Math Vocabulary raw score
22. Reading Speed raw score
23. Reading Comprehension raw score
24. Reading Efficiency raw score

Table entries are correlation coefficients corrected for attenuation, the number of cases upon which they are based, and the uncorrected correlation. Only correlations significant at $p < .01$ are presented. Decimals have been omitted. Pairwise deletion was used in the calculation of the multiple correlations, which are not corrected for attenuation.

[a]The disattenuated correlations for the Global-Analytic style scale are omitted because of the very low reliability of that scale (.03)

measures. The simple (individual) correlations reported in the table were corrected for attenuation (unreliability). The simple correlations used to compute the multiple correlation were not corrected for attenuation; because of this, in some cases the simple correlations were higher than the multiple correlation for the same variable.

Ability score on the Global-Analytic test was correlated at a high (> .6) level with Analytical Reasoning, Wiggly Block, Paper Folding, Math Vocabulary, and Number Series. Global-Analytic ability was also moderately correlated with Memory for Design (.50) and English Vocabulary (.49). The correlations between Global-Analytic ability and some of these measures may be due to their shared assessment of a general reasoning factor. The multiple correlation between Global-Analytic ability and the standard battery variables was .70, indicating that Global-Analytic ability overlaps substantially with the measures already in the JOCRF battery.

Scores on Object Assembly were correlated moderately highly ($r > .58$) with Wiggly Block, Paper Folding, Analytical Reasoning, and Memory for Design. Object Assembly, then, did predict spatial aptitude as hypothesized. The multiple correlation was .56, indicating that Object Assembly scores are moderately well predicted from battery measures. MFFT-ability did not correlate above .26 with any battery variable, and its multiple correlation was only .31.

RFT scores did not correlate above .32 with any battery variable. The strongest relationship with the RFT was found for a spatial aptitude measure, Paper Folding. The multiple correlation for the RFT was .35, indicating only a modest degree of overlap with the battery tests. Battery variable correlations with CWR did not reach a moderate level for any variable. MFFT-style correlated .35 with Reading Comprehension. All other correlations for MFFT-style were less than .30. The correlations for Global-Analytic style were all low or not significant. Its highest correlation was with Math Vocabulary (.21 without disattenuation). Scores on the VVQ were moderately correlated with English Vocabulary (-.59), Reading Comprehension (-.59), Reading Speed (-.36), and Reading Efficiency (-.53). VVQ scores were not related above a low level with spatial ability measures, as had been hypothesized. The VVQ was moderately well-predicted by the battery variables, with a multiple correlation of .49.

Factor Structure of Battery Variables and Cognitive Style

Principal components analyses with varimax rotation were performed with different sets of battery variables. The first analysis did not include all battery variables, nor did it include Global-Analytic ability nor MFFT-ability. This analysis was performed with both listwise and pairwise deletion of cases with missing values. Listwise deletion resulted in exclusion of the majority of cases, and so the results reported are from the analysis using pairwise deletion.

When 14 battery variables and six cognitive style variables were used, six factors with eigenvalues greater than 1.0 were found, accounting for 57% of the variance. Table 5 presents the variables and factor loadings for these six factors. MFFT-style and Memory for Design loaded on more than one factor. RFT loaded on the factor with spatial aptitude measures, as hypothesized.

Using 24 battery variables, including vocabulary and reading measures, along with six cognitive style measures, nine factors with eigenvalues greater than 1.0 were found, accounting for 61% of the variance. Table 6 presents the variables and factor loadings for these nine factors. Again, pairwise deletion was used.

These principal components analyses suggest that CWR and Global-Analytic style form a factor independent of battery measures, but that there is some overlap between VVQ, RFT, and Object Assembly, and battery variables. There is also some overlap between MFFT-style and battery variables, but the loadings of MFFT-style are split and low or negative. The zero-order correlations and multiple regressions indicate that the overlap for RFT and MFFT-style is low.

Sex, Age, Education, and Laterality Differences

Significant sex differences ($p < .01$, were found for RFT, VVQ, and Global-Analytic ability (see Table 7). Women were more field dependent than men, as was hypothesized. Women made more-severe errors than men in terms of undercorrection (deviation from the vertical in the direction of the frame; $p < .01$), but not in overcorrection (deviation from the vertical away from the direction of the frame; $p > .01$). This is consistent with findings in the research literature, which indicate that women are more likely to align the rod in the direction of the

Table 5

Principal Components Analysis of Battery and Cognitive Style Variables (Reduced Set)

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Paper Folding	.77					
Wiggly Block	.76					
Memory for Design	.54	.54				
Analytical Reasoning	.50					
Number Series	44					
Object Assembly	.68					
MFFT-style	.39					-.33
RFT	.53					
Number Memory		.79				
Silograms		.74				
Observation		.68				
Tonal Memory			.83			
Pitch Discrimination			.75			
Rhythm Memory			.69			
Ideaphoria				.81		
Foresight				.77		
Inductive Reasoning				.44		
Global-Analytic style					.81	
CWR					.58	
VVQ						.74

frame than are men. On the VVQ, women scored in the direction of a more visual, and less verbal, style than men, which contrasts with the customary sex-difference findings of female superiority on verbal fluency and male superiority on spatial ability (Maccoby & Jacklin, 1974; Technical Report 1986-1). Men outperformed women on the Global-Analytic test, with significantly higher ability scores. Global-Analytic style differences approached significance ($p < .02$), with men being less analytically and more globally oriented. MFFT-ability differences also approached significance ($p < .03$), with women performing at a higher level than men. Sex differences for MFFT-style and Object Assembly scores were not significant.

Sex differences in color, word, and color-word naming were all significant at $p < .01$, with females outperforming males on

Table 6

Principal Components Analysis of Battery and Cognitive Style Variables (Complete Set)

Variable	Factor 1	2	3	4	5	6	7	8	9
Paper Folding	.75								
Wiggly Block	.78								
Memory for Design	.56	.52							
Analytical Reasoning	.50		.32						
Object Assembly	.69								
RFT	.47								.30
Number Memory		.79							
Silograms		.70							
Observation		.65							
Graphoria		.49							
Number Series	.40	.41	.39						
English Vocabulary			.79						
Reading Comprehension	.32	.39	.71						
VVQ			-.72						
Ideaphoria				.83					
Writing Speed				.74					
Foresight				.65					
Tonal Memory					.83				
Pitch Discrimination					.75				
Rhythm Memory					.69				
Reading Speed						.85			
Reading Efficiency			.61			-.69			
Global-Analytic style						.84			
CWR							.46	.25	-.33
Tweezer Dexterity								.59	
Finger Dexterity				.39				.49	
Color Perception								.65	
Personality									.70
Inductive Reasoning	.40								.49
MFFT-style									-.45

Table 7

Sex Differences in Cognitive Style

Measure	Female Mean	Female SD	Male Mean	Male SD	t	p	Effect size
Rod-and-Frame Test raw score	21.18 285	22.84	15.21 269	13.48	3.77	.01	.28
Undercorrection	1.97 291	2.66	1.34 269	1.62	3.09	.01	.25
Overcorrection	.61 291	.88	.48 269	.49	2.24	.03	.19
Color-word residual raw score	1.10 488	7.68	.44 437	7.01	1.37	>.05	.09
MFFT-ability	.35 453	1.00	.22 417	.77	2.20	.03	.15
MFFT-style	-.21 453	1.80	-.17 417	1.62	-.34	>.05	.02
Verbalizer-Visualizer	6.41 440	2.00	5.87 398	2.18	3.77	.01	.26
Object Assembly	3.01 475	1.00	3.02 433	.97	.05	>.05	.01
Global-Analytic -ability	-.31 294	1.51	.30 257	1.84	4.19	.01	.36
Global-Analytic -style	.11 294	1.00	-.09 257	1.09	2.33	.02	.19

Note. Effect size is the result of the difference between groups divided by the overall sample standard deviation. The number of cases in each group is provided beneath the score means.

each task. (The means for females and males were: words--M_f = 109.5, M_m = 104.6; colors--M_f = 80.2, M_m = 75.2; color-words--M_f = 47.2, M_m = 44.0.) Differences favoring females have been found for colors and color-words in past research (Golden, 1978). The present study employed a larger sample size than that of most previous research efforts, which may have resulted in the significant sex difference in word scores. Consistent with past research, sex differences in CWR scores were not significant.

Age effects were assessed using one-way weighted-means analyses of variance. Age effects were found for four of the eight cognitive style variables: CWR, MFFT-style, VVQ, and Global-Analytic style (see Table 8). CWR scores decreased with age, indicating greater susceptibility to distraction with age. Cognitive style became less impulsive and more reflective with age. Visual-verbal preference tended to be more verbal and less visual with age as well. Cognitive style became more global and less analytic as age increased.

The relationship between educational level and cognitive style was assessed using one-way weighted-means analyses of variance. Significant relationships were found for CWR, MFFT-style, VVQ, and Global-Analytic ability (see Table 9). Distractibility and reflectivity increased as educational level increased, and examinees scored more in the verbal and less in the visual direction on the VVQ. Ability on the Global-Analytic test also increased with educational level.

Laterality differences were assessed using _t_-tests and one-way weighted-means analyses of variance. No significant relationships were found for any measure (see Table 10). This is in contrast to expectations for RFT performance. Right-handed persons and those with a strong eye preference (left or right) were expected to be more field independent but were not found to be so.

Cognitive Style Differences by Major

Differences among college majors were assessed using one-way weighted-means analyses of variance. Majors were categorized to provide a minimum of 15 persons per major in the sample as a whole. Further grouping was needed with cognitive style variables with fewer total cases. This provided the following groups:

1. Agriculture
2. Art
3. Biological Sciences
4. Business
5. Communication
6. Computer Science
7. Education
8. Engineering
9. English
10. Foreign Languages
11. Health and Medicine
12. History
13. Music
14. Physical Sciences
15. Psychology
16. Social Sciences

Table 8

Age Differences in Cognitive Style

Variable	Age 14-18	19-23	24-34	35+	F	p	Effect size	r
Rod-and-Frame Test	.11	.21	.02	-.13				
_z_-score	(1.10)	(.60)	(.86)	(.97)	3.60	.02	.38	-.12
	113	119	157	165				554
Color-word residual	.30	.14	-.02	-.37				
_z_-score	(.92)	(.96)	(1.07)	(1.03)	18.95	.01	.67	-.23
	222	231	236	236				925
MFFT-ability	.38	.34	.26	.18				
	(1.20)	(.79)	(.77)	(.75)	2.32	.07	.23	-.09
	210	220	225	215				870
MFFT-style	-.52	-.42	-.02	.19				
	(2.10)	(1.64)	(1.61)	(1.34)	8.51	.01	.42	.16
	210	220	225	215				870
Verbalizer-	6.52	6.50	5.89	5.65				
Visualizer	(1.86)	(2.07)	(2.20)	(2.17)	9.26	.01	.41	-.20
	217	214	208	199				838
Object Assembly	2.99	3.10	3.02	2.95				
	(.98)	(.94)	(1.02)	(.99)	1.04	.37	.15	-
	221	229	232	226				
Global-Analytic	-.07	-.12	-.04	.12				
-ability	(1.64)	(1.70)	(1.65)	(1.81)	.53	.66	.14	.05
	137	133	137	144				551
Global-Analytic	.28	.14	.04	-.37				
-style	(.96)	(.93)	(1.04)	(1.12)	10.50	.01	.62	-.20
	137	133	137	144				551

Note. Effect size is the result of the largest difference between groups divided by the overall sample standard deviation. Standard deviations are in parentheses, with sample sizes on the line below.

Table 9

Educational Level Differences in Cognitive Style

Variable	0-12	13-15	16	17-22	F	p	Effect size	r
Rod-and-Frame Test z-score	.06 (1.08) 137	.12 (.72) 165	-.03 (.97) 142	-.04 (.84) 110	1.05	.37	.16	-
Color-word residual z-score	.22 (.87) 264	.01 (1.05) 290	-.12 (1.01) 213	-.18 (1.17) 158	7.12	.01	.40	-.16 913
MFFT-ability	.32 (1.17) 250	.27 (.77) 272	.23 (.74) 199	.34 (.79) 149	.57	.63	.12	-
MFFT-style	-.61 (2.04) 250	-.21 (1.58) 272	-.01 (1.47) 199	.31 (1.49) 149	10.37	.01	.53	.17 858
Verbalizer-Visualizer	6.48 (1.93) 248	6.39 (2.11) 275	5.96 (2.11) 187	5.30 (2.17) 128	11.03	.01	.56	-.20 829
Object Assembly	2.90 (1.03) 253	3.06 (.95) 278	3.07 (.99) 206	3.05 (.93) 149	1.67	.17	.18	-
Global-Analytic -ability	-.23 (1.66) 161	-.16 (1.63) 180	.07 (1.76) 124	.51 (1.75) 86	4.18	.01	.44	.15 544
Global-Analytic -style	.19 (.91) 161	-.02 (1.02) 180	-.07 (1.24) 124	-.10 (1.00) 86	2.18	.09	.28	-.13 544

Note. Effect size is the ratio of the largest difference between groups divided by the overall sample standard deviation. Standard deviations are in parentheses with sample sizes on the line below.

Table 10

Laterality Differences in Cognitive Style

Variable	RH/RE	LH/LE	LH/RE	RH/LE	Other	F	p
Sample size	340	30	32	139	522		
Rod-and-Frame Test z-score	.02	.02	.04	-.03	-.01	.06	.99
Color-word residual z-score	.03	.02	.11	-.00	-.02	.23	.92
MFFT-ability	.29	.13	.39	.32	.30	.34	.85
MFFT-style	-.21	-.11	-.43	-.17	-.22	.16	.96
Verbalizer-Visualizer	6.29	6.43	5.47	6.28	6.06	1.63	.16
Object Assembly	2.94	3.06	3.08	3.08	3.06	.89	.47
Global-Analytic -ability	-.01	-.27	.85	-.23	.01	1.84	.12
Global-Analytic -style	.11	.09	-.16	.02	-.08	1.15	.33

(cont.)

Table 10 (cont.)

Variable	RH	LH	t	p	RE	LE	t	p
Rod-and-Frame Test z-score	.02 388	.07 53	.46	.64	.01 267	-.01 136	.23	.82
Color-word residual z-score	-.04 647	.04 84	.65	.52	.03 453	.00 224	.42	.67
MFFT-ability	.29 609	.36 77	.71	.48	.28 424	.27 208	.18	.85
MFFT-style	-.14 609	-.20 77	.29	.78	-.22 424	-.14 208	.58	.56
Verbalizer-Visualizer	6.27 592	5.72 75	2.11	.04	6.20 403	6.29 201	.53	.60
Object Assembly	3.00 636	3.16 83	1.39	.16	2.95 441	3.10 216	1.77	.08
Global-Analytic -ability	-.01 381	.25 51	1.02	.31	.02 271	-.28 136	1.72	.08
Global-Analytic -style	.04 381	.02 51	.10	.92	.04 271	.06 136	.19	.85

Note. Left-handed (LH) is a score of 0-.50 on a scale of 0-1.0; right-handed (RH) is a value of 1.0 on this scale.

RFT performance was the most field dependent for the health and medicine majors and the most independent for the physical sciences majors (diff/std dev = .50). The health and medicine category was composed of dentists/dental assistants, doctors, nurses, occupational and physical therapists, and so on. Significant differences were found between health and medicine majors and business, engineering, and psychology majors ($p < .05$). Results of previous studies suggest persons in socially oriented fields to be more field dependent than persons in analytically oriented fields. In the present study, differences were not found for education majors, typically considered to be a socially oriented group, but were found for health and medicine majors, also considered to be a socially oriented group (with some exceptions). Interpretation of these results as consistent with previous findings depends on how majors are categorized as socially vs. analytically oriented. Since the groups were small in size, it is suggested that less credence be placed in this analysis than in analyses by sex, age, or education.

Majors were further grouped as art/music, biological sciences/health and medicine, business, education, and social science majors to provide groups with at least 20 cases. Significant overall differences ($F_{4,108} = 3.83$, $p < .01$) were found, with biological science/medicine majors being more field dependent than business majors.

Overall differences in ability scores on the Global-Analytic test were significant ($F_{4,185} = 2.83$, $p < .03$), but no two groups were significantly different when the collapsed coding was used. Education majors had the lowest scores ($M = -.46$, $n = 28$), and social science majors the highest ($M = .57$, $n = 29$). When the original grouping of 16 majors listed above was used, overall differences were significant, and computer science majors were found to be significantly more able than history majors.

No significant differences among majors were found for CWR, MFFT-ability, MFFT-style, VVQ, Object Assembly, or Global-Analytic style scores.

GENERAL SUMMARY AND DISCUSSION

The RFT had very good internal consistency reliability (.93) and was correlated at a low level with other cognitive style variables. Correlation with battery variables was highest with spatial aptitude measures but was still low. Significant age and sex differences were found in RFT scores but not the expected differences in laterality.

The reliability of CWR scores could not be assessed in this study and will need to be assessed if further research on the test is to be done. Test-retest, split-half, or parallel forms reliability could be estimated in the future, given the appropriate test development and administration. CWR scores were somewhat correlated with Global-Analytic style, with a correlation of .15 uncorrected for unreliability. (Because of the extremely low reliability of Global-Analytic style, it is not possible to estimate the corrected correlation.) CWR scores were correlated at a low level with battery variables and formed a separate factor in conjunction with Global-Analytic style. No sex or laterality differences were found, but there were significant effects of age and educational level on CWR scores, with resistance to distractibility decreasing with greater age and educational level.

The internal consistency reliability of the MFFT-style scale was above the JOCRF recommended minimum standard of .80 (Test Information Bulletin 1980-7). The internal consistency reliability of the ability score was substantially lower, due to the negative relationship between accuracy and speed and the low reliability of the accuracy score. Additional items or revised items could be used to increase MFFT accuracy reliability, which could be expected to increase the reliability of the ability score. Neither the ability nor the style score was correlated above a moderate level with other cognitive style or battery variables. Significant effects for age and educational level were found for style, with style being more reflective with increased age and educational level.

The internal consistency reliability of VVQ scores was low, and subsequent analysis indicated that the items were not measuring a unitary bipolar trait. Items were of a limited range of difficulties, and it seemed that they were measuring two distinct variables: (a) liking for visual thought or activities and (b) liking for verbal thought or activities. When two separate subscales were formed, the verbal subscale correlated at a moderate to high level with vocabulary and reading variables, but the visual subscale did not correlate beyond a low level with spatial aptitude measures. Either visual preference does not relate to spatial aptitude, or the measurement of visual preference is suspect. The item set, then, forms the core for development of two measures that would bear further investigation. Correlations for overall VVQ with other cognitive style measures (except for Global-Analytic ability) were low, as were correlations with all battery variables except verbal performance measures (vocabulary, reading measures). Significant differences were found for sex, age, and educational level but not for laterality.

The internal consistency reliability of Object Assembly was low. Object Assembly was included in this study as a potential validation measure rather than as a potential cognitive style measure. Among the experimental measures, Object Assembly had the highest correlations with other cognitive style measures and with battery variables. Object Assembly was related neither to Global-Analytic style nor to reflectivity-impulsivity, as had been hypothesized. It was correlated with other spatial aptitude measures, however. No significant effects of sex, age, educational level, or laterality were found.

The most tenuous results were found for the Global-Analytic scales (ability and style). The reliability of the ability scale was .72, and the scale correlated with measures of reasoning and other aptitudes. The reliability of the style scale was extremely low, and so relationships between that variable and others are unclear. Analyses suggest that global ability and analytic ability are positively related, not bipolar. The Global-Analytic ability scale may be useful as a proxy for reasoning and/or spatial aptitude measures. The style scale, on the other hand, is of dubious value. The effort to develop a general, bipolar measure of global-analytic style was unsuccessful.

The purpose of this project was to assess the psychometric value of a number of experimental measures, the relationships among them, and their relationships to the aptitudes already assessed by the standard JOCRF battery. The reliability values for the RFT and MFFT-style scales were acceptable, while reliability values for the remaining tests indicate that some revision will be necessary to provide adequate measurement precision. The distractibility and global-analytic ability measures showed close-to-adequate reliability (.70 and above) and with further refinement could be acceptable. It should be noted that reliability for the CWR scores was not calculated in this study but has been estimated at .70.

The cognitive style measures were selected to assess different dimensions of cognitive style. Correlational and factor analyses indicated that, with the exception of Object Assembly and Global-Analytic ability, measures were not correlated even at a moderate level with each other. And prediction from battery variables was low, further indicating that the cognitive style measures were assessing constructs different from those measured by battery variables. This is particularly true for Global-Analytic style and for the CWR.

The results of this study suggest that the confounding of styles and abilities criticized by Shipman and Shipman (1985) can be minimized and that for some style measures, adequate reliability can be achieved. This study did not address the cross-task generality of styles, an aspect of cognitive style that needs to be investigated. Whether these constructs are important for vocational guidance and whether they are related to real-world behaviors remain to be investigated.

REFERENCES

Boyden, J. G., & Gilpin, A. R. (1978). Matching Familiar Figures Test and Stroop test performance in adults. Perceptual and Motor Skills, 46, 854.

Golden, C. J. (1978). Stroop Color and Word Test (test manual). Chicago: Stoelting.

Heckel, R. V., Hiers, J. M., Laval, C. J., & Allen, S. S. (1980). Adult norms on the Kagan Matching Familiar Figures Test of impulsivity/reflectivity. JSAS Catalog of Selected Documents in Psychology, 11, 5.

Kagan, J., Pearson, L., & Welch, L. (1966). Conceptual impulsivity and inductive reasoning. Child Development, 37, 359-365.

Jensen, A. R., & Rohwer, W. D., Jr. (1966). The Stroop Color-Word Test: A review. Acta Psychologica, 25, 36-93.

Maccoby, E. E., & Jacklin, C. N. (1974). The psychology of sex differences. Stanford, CA: Stanford University Press.

Messer, S. B. (1970). Reflection-impulsivity: Stability and school failure. Journal of Educational Psychology, 61, 487-490.

Messer, S. B. (1976). Reflection-impulsivity: A review. Psychological Bulletin, 83, 1026-1052.

Minutes of research committee meeting. (February, 1975). M. Daniel, unpublished document.

Oltman, P. K. (1968). A portable rod-and-frame apparatus. Perceptual and Motor Skills, 26, 503-506.

Paivio, A. (1971). Imagery and verbal processes. New York: Holt, Rinehart.

Richardson, A. (1977). Verbalizer-visualizer: A cognitive style dimension. Journal of Mental Imagery, 1, 109-126.

Shipman, S., & Shipman, V. C. (1985). Cognitive styles: Some conceptual, methodological and applied issues. In E. W. Gordon (Ed.), Review of research in education, 12 (pp. 229-291). Washington, DC: American Educational Research Association.

Spoltore, J. D., & Smock, D. J. (1983). The Verbalizer-Visualizer Questionnaire: Additional normative data. Perceptual and Motor Skills, 56, 382.

SPSS Inc. (1986). SPSS-X user's guide (2nd ed.). New York: McGraw-Hill.

Stroop, J. R. (1935). The basis of Ligon's theory. American Journal of Psychology, 47, 499-504.

Technical Report 1985-1. Cognitive style: A review of the literature. K. Green. Chicago: Johnson O'Connor Research Foundation.

Technical Report 1986-1. The measurement of human variation in spatial visualizing ability: A process-oriented perspective. M. F. Zimowski & W. Wothke. Chicago: Johnson O'Connor Research Foundation.

Test Information Bulletin 1980-7. Reliabilities and standard errors of laboratory worksamples. M. Daniel. New York: Johnson O'Connor Research Foundation.

Wechsler, D. (1981). Wechsler Adult Intelligence Scale-Revised. New York: Psychological Corporation.

Wright, B. D., Mead, R. J., & Bell, S. R. (1980). BICAL: Calibrating items with the Rasch model (Research Memorandum 23C). Chicago: University of Chicago, Statistical Laboratory, Department of Education.

APPENDIX A

CALCULATION ROUTINE FOR OVER- AND UNDER-CORRECTION ON THE ROD-AND-FRAME TEST

```
RECODE          Dir2, Dir3, Dir6, Dir7 (2=1)(1=2)
DO REPEAT       O=O1 to O8/U=U1 to U8/D=Dir1 to Dir8/
                  T=Trial1 to Trial8
IF              (D eq 1) U=T
IF              (D eq 2 or D eq 3) U=0
IF              (D eq 2) O=T
IF              (D eq 1 or D eq 3) O=0
END REPEAT
COMPUTE         Under=Mean(U1,...,U8)
COMPUTE         Over=Mean(O1,...,O8)
```

APPENDIX B

ITEM STATISTICS

Rod-and-Frame Test

Measure	Mean	SD	N	Kurtosis	Skewness	Range	Corr
Trial 1	2.92	3.68	650	11.5		0,28	.71
Trial 2	2.61	3.33	650	18.9	3.5	0,29	.76
Trial 3	2.15	3.00	647	29.7	4.5	0,28	.76
Trial 4	1.91	2.91	650	32.1	4.9	0,27	.78
Trial 5	2.51	3.34	651	21.9	3.9	0,28	.81
Trial 6	2.56	3.25	648	17.4	3.5	0,26	.77
Trial 7	2.16	2.92	649	30.4	4.5	0,28	.78
Trial 8	2.11	2.89	649	28.7	4.5	0,27	.78

Trials were: 1--F28 left, R28 left; 2--F28 left, R28 right; 3--F28 right, R28 right; 4--F28 right, R28 left; 5--F28 left, R28 left; 6--F28 left, R28 right; 7--F28 right, R28 right; 8--F28 right R28 left. Read as Frame 28 degrees left or right and Rod 28 degrees left or right.

Stroop Color and Word Test

Measure	Mean	SD	N	Kurtosis	Skewness	Range
Words	107.21	15.47	1079	1.6	-.2	14,155
Word errors	.85	1.23	1077	8.0	2.3	0,10
Colors	77.87	11.34	1076	.5	-.2	22,115
Color errors	1.37	1.49	1076	3.1	1.5	0,9
Color-words	45.68	8.67	1078	.7	.0	4,81
Color-word errors	1.59	2.01	1078	10.7	2.5	0,18
Color-word predicted score	44.92	6.06	1076	1.2	-.3	11.98, 62.73

Matching Familiar Figures Test

	Item	Mean Time	SD	N	Kurtosis	Skewness	Range	Corr
1.	Dog	74.66	60.68	1081	21.1	3.2	2,675	.65
2.	Flower	94.74	69.83	1081	5.7	1.6	5,640	.74
3.	Soldier	112.33	78.36	1081	2.6	1.3	5,599	.75
4.	Graph	88.24	61.62	1081	13.3	2.5	6,704	.73
5.	Child	127.69	93.82	1081	6.0	1.8	6,821	.77
6.	Lamp	49.62	35.67	1080	41.2	4.8	2,504	.62
7.	Dress	114.09	81.75	1072	6.1	1.8	3,705	.78
8.	Lion	135.00	90.30	1072	2.4	1.1	4,710	.79
9.	Sunglasses	65.95	41.21	1070	12.0	2.4	4,407	.59
10.	Airplane	90.01	51.07	1066	2.2	1.1	4,389	.72
11.	Leaf	87.04	51.54	1066	3.1	1.3	4,370	.68
12.	Bed	68.14	42.36	1066	7.6	2.0	2,418	.55

	Item	% Correct	N	Corr
1.	Dog	48	1081	.26
2.	Flower	39	1081	.37
3.	Soldier	59	1081	.29
4.	Graph	61	1081	.33
5.	Child	56	1080	.38
6.	Lamp	74	1080	.48
7.	Dress	53	1072	.34
8.	Lion	43	1072	.41
9.	Sunglasses	74	1070	.38
10.	Airplane	70	1067	.35
11.	Leaf	75	1067	.27
12.	Bed	76	1058	.44

Verbalizer-Visualizer Questionnaire

Item	% true	N	Corr
1. I enjoy doing work that requires the use of words. (R)	75	981	.25
2. My daydreams are sometimes so vivid I feel as though I actually experience the scene.	65	981	.32
3. I enjoy learning new words. (R)	84	981	.19
4. I can easily think of synonyms for words. (R)	58	981	.28
5. My powers of imagination are higher than average.	67	981	-[a]
6. I seldom dream. (R)	17	981	.20
7. I read rather slowly.	49	981	.15
8. I cannot generate a mental picture of a friend's face when I close my eyes. (R)	9	981	-
9. I don't believe that anyone can think in terms of mental pictures. (R)	2	981	-
10. I prefer to read instructions about how to do something rather than have someone show me. (R)	29	981	.12
11. My dreams are extremely vivid.	68	981	.21
12. I have better than average fluency in using words. (R)	57	981	.34
13. My daydreams are rather indistinct and hazy. (R)	25	981	.30
14. I spend very little time attempting to increase my vocabulary.	59	981	-
15. My thinking often consists of mental pictures or images.	84	981	.26

Note. (R) = coding was reversed for this item.

[a] Item dropped from scale.

Object Assembly

Item	Mean Accuracy	SD	N	Kurtosis	Skewness	Range	Corr
Item 1: Manikin	4.97	.34	1087	82.4	-5.8	1,8	.32
Item 2: Profile	8.68	.98	1082	21.8	-4.2	0,9	.44
Item 3: Hand	6.68	94	1079	16.9	-3.6	0,9	.32
Item 4: Elephant	7.43	1.96	1056	8.5	-3.2	0,9	.34

Item	Mean Time	SD	N	Kurtosis	Skewness	Range	Corr
Item 1: Manikin	18.96	9.14	1087	27.0	3.8	2,120	.33
Item 2: Profile	55.94	28.37	1086	.1	1.0	1,120	.41
Item 3: Hand	78.23	41.22	1082	.5	1.1	1,180	.33
Item 4: Elephant	86.32	58.94	1064	-1.2	.5	1,180	.32

Global-Analytic Problem-Solving Test

Item	Mean	SD	N	Kurtosis	Skewness	Corr
Analytic						
Letters	1.14	1.16	604	-1.3	.4	.37
Anagrams	1.27	1.21	605	-1.5	.3	.27
River	.55	1.31	605	1.0	1.6	.39
Barnyard	.83	1.05	588	.2	.7	.34
Clock	.50	.96	586	-1.1	.4	.25
Logic	1.75	1.04	605	-1.1	-.3	.28
Global						
Horse	1.60	1.12	605	-1.3	-.2	_[a]
Hunter	.44	.93	605	2.3	1.9	.31
Ship	.45	1.00	605	2.1	2.0	.34
Points	.79	1.01	605	-.1	1.1	.33
Coins	1.06	1.23	605	-1.4	.6	.20
Twos	1.64	1.21	605	-1.5	-.3	.23
Sayings: 1	1.05	1.21	604	-1.4	.5	_[b]
2	.59	1.10	604	.4	1.5	_[b]
3	1.00	1.24	604	-1.3	.7	_[b]
Sayings combined	.88	.77	604	-.4	.6	.26

Note. Each of the items on the Global-Analytic test was scored on a four-point scale, from 0 to 3, with the score based on the examinee's time to correct solution (see Table 2.). Examinees who answered an item incorrectly were given a score of 0 on that item.

[a]Dropped from scale.

[b]Used only as a summed score and not individually.

APPENDIX C

DESCRIPTIVE STATISTICS FOR STANDARD BATTERY MEASURES

Measure	Mean	SD	N	Kurtosis	Skewness	Range
Age	27.32	10.41	972	-.2	.8	14,65
Education	14.28	2.66	960	-.5	.1	8,22
(Right-) Eyedness	.65	.43	967	-1.4	-.6	0,1.0
(Right-) Handedness	.90	.23	971	6.4	-2.8	0,1.0
Graphoria percentile score	58.46	27.79	972	-1.1	-.3	0,99
Color Perception RS	13.40	1.87	964	23.2	-4.7	0,14
Ideaphoria RS	277.30	69.30	965	.5	.1	35,527
Foresight RS	46.53	19.31	969	.7	1.0	9,119
Inductive Reas. RS	143.75	21.95	958	-.1	-.2	59,198
Analyt. Reas. RS	33.13	6.78	930	-.3	-.3	14,48
Wiggly Block RS	249.58	103.47	966	-.9	.1	5,488
Paper Folding RS	19.68	12.89	968	-.2	.7	0,60
Personality RS	16.33	8.27	972	-.8	-.2	0,38
Tonal Memory RS	56.66	14.24	972	-.1	-.5	0,85
Pitch Discrim. RS	63.18	9.66	972	.4	-.8	20,80
Rhythm Memory RS	45.57	6.13	972	14.8	-2.6	0,55
Memory for Design RS	79.31	26.11	971	.0	.1	0,160
Silograms RS	19.52	9.55	964	-.9	.2	0,40
Number Memory RS	80.66	28.41	966	-.4	-.1	0,144
Observation RS	67.14	11.42	873	-.2	-.1	36,102
Tweezer Dexterity RS	42.79	18.31	952	-1.0	-.1	5,80
Finger Dexterity RS	75.75	12.34	954	-.1	.1	35,119
Number Series RS	23.13	4.74	717	1.9	-1.1	1,30
English Vocabulary Scale Score	148.03	39.05	972	-.9	-.2	57,222
Math Vocabulary RS	26.28	8.54	532	-.6	.2	6,48
Reading Speed RS	246.58	85.05	224	1.4	.9	65,615
Reading Comp. RS	23.62	6.71	225	-.8	-.3	7,36
Reading Effic. RS	25.06	6.48	706	.1	-.3	2,40

JOHNSON O'CONNOR RESEARCH FOUNDATION

HUMAN ENGINEERING LABORATORY