7-1989

The Effects of Adding Verbalizations on the Draw-A-Person

Phyllis Helen Millspaugh

Follow this and additional works at: https://digitalcommons.wku.edu/theses
Part of the Psychology Commons

Recommended Citation
https://digitalcommons.wku.edu/theses/2062
THE EFFECTS OF ADDING VERBALIZATIONS
ON THE DRAW-A-PERSON

A Thesis
Presented to
the Faculty of the Department of Psychology
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
of the requirements for the Degree
Master of Arts

by
Phyllis Helen Millspaugh
July 1989
AUTHORIZATION FOR USE OF THESIS

Permission is hereby

☒ granted to the Western Kentucky University Library to make, or allow to be made photocopies, microfilm or other copies of this thesis for appropriate research for scholarly purposes.

☐ reserved to the author for the making of any copies of this thesis except for brief sections for research or scholarly purposes.

Signed: [Signature]

Date: 7/12/89

Please place an "X" in the appropriate box.

This form will be filed with the original of the thesis and will control future use of the thesis.
THE EFFECTS OF ADDING VERBALIZATIONS ON THE DRAW-A-PERSON

Recommended 7/10/89  
[Signature]  
Director of Thesis

[Signature]  
Lynn Clark

Approved July 12, 1989  
[Signature]  
Dean of Graduate College
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the members of my committee for their assistance and support during this project. Dr. Sally Kuhlenschmidt was not only instrumental in guiding me through this novel experience, providing critical thinking skills when mine failed, and teaching organization skills that were previously foreign to me, but she was also a firm, but flexible sounding board that allowed many sessions of frustration to be heard and expelled. Dr. Daniel Roenker played a vital role in this project by providing his infinite statistical knowledge, unflappable sense of logic, and enduring sense of humor to an otherwise painful year of researching. Dr. Lynn Clark was of tremendous value by critically and creatively analyzing my writing throughout this project.

Additionally, a special note of thanks belongs to Michela LaRocca, Rodney Kinslow, Kenneth Hinton, Cindi Eaton, and Tim Robins. Without the willingness, enthusiasm, responsibility, and intellectual curiosity displayed by these five research assistants, this project could not have been completed. They were an invaluable part of the data collection, and I am greatly indebted to them for their effort.

I owe a big thank-you to Aaron Peters for his consistent attempts at making my sessions at the computer terminal easy ones. Any research project requires ample statistical analyses that can quickly become frustrating experiences. Through Aaron’s good nature and technical knowledge, I was able to master the demon inside the computer long enough to allow me to finish my research.

Finally, I would like to express my thanks to Cindy Sellers, who reviewed and critiqued my writing throughout the year of research. Without her expert grammatical and stylistic skills, I could not have written a paper that made not only statistical sense, but logical sense. I will forever remember, "Don’t make the reader work."
Table of Contents

ACKNOWLEDGEMENTS ......................................................... iii
LIST OF TABLES .............................................................. iii
LIST OF APPENDICES ....................................................... v
ABSTRACT ................................................................. vi
INTRODUCTION ............................................................. vii
DRAWING TEST BACKGROUND ........................................... 1
PSYCHOMETRIC PROPERTIES/DIFFICULTIES .............................. 2
HYPOTHESES .............................................................. 7
METHOD ................................................................. 9
RESULTS ................................................................. 11
DISCUSSION .............................................................. 16
APPENDIX A .............................................................. 19
APPENDIX B .............................................................. 24
APPENDIX C .............................................................. 25
APPENDIX D .............................................................. 26
APPENDIX E .............................................................. 29
APPENDIX F .............................................................. 30
APPENDIX G .............................................................. 32
REFERENCES ............................................................. 34

iv
List of Tables

Table                                      Page

I  Means and Standard Deviations by Grade Level
   for Male and Female Draw-a-Person Total Scores                           17

II Correlation Matrix for Verbal and Nonverbal
   Groups With All Tests                                                   18
THE EFFECTS OF ADDING VERBALIZATION
THE DRAW-A-PERSON

Phyllis Helen Millspaugh
July 1989

Directed by: Sally L. Kuhlenschmidt, Daniel Roenker, and Lynn Clark
Department of Psychology
Western Kentucky University

Sixty-two male and female 9-10-, and 11-year olds were administered the Naglieri
Draw-A-Person (DAP) and the Vocabulary and Block Design subtests from the Wechsler
Intelligence Scale for Children-Revised (WISC-R). The study was conducted to determine the
effects of verbalizations during the DAP on correlations with an estimated Full Scale IQ Score
obtained from the two WISC-R subtests. Results of the study indicate that although there was
an increase in the correlations between the Verbal DAP group and the estimated Full Scale IQ
Score, compared to the standard DAP and the Full Scale IQ, it was not significant. The
difference between the mean scores on the Verbal DAP group and the Nonverbal DAP group
was also nonsignificant. The overall correlations between the DAP and the estimated Full Scale
IQ Scores showed no improvement from previous attempts to correlate standard intelligence
tests with drawing tests. Further research is suggested to determine the effects of a larger
sample size on a similar design.
The Effects of Adding Verbalizations on the Draw-a-Person

A popular theory of intelligence (Wechsler, 1974) holds that global intelligence is made up of two components: verbal and perceptual motor, or performance, abilities. Many standardized intelligence tests, like the Wechsler intelligence scales, are based on this concept of a general ability or combination of verbal and performance abilities. An alternative and popular method of measuring intelligence is the human figure drawing test, apparently a spatial or performance measure. This test is popular because it is a quick assessment. Unfortunately, it appears to lack a verbal component. The current study investigates if and how adding verbalization effects the validity of the Draw-a-Person (DAP). A quick screening test for intelligence could be of great utility for the researcher. A brief intelligence test could potentially help a school psychologist narrow down a child's weaknesses and focus in on the major difficulties more quickly; it could also prevent a child from missing valuable classroom instruction time. With these potential benefits in mind, an evaluation of the drawing test's use as a quick screening test for intelligence is in order. Specifically, by comparing IQ scores obtained from the DAP to those obtained from the Vocabulary and Block Design subtests of the Wechsler Intelligence Scale for Children-Revised (WISC-R)--selected because they are the purest measures of their respective scales on the WISC-R--the utility of the DAP as a valid, quick-screening device can be evaluated. An additional area of interest is whether or not the correlations between the DAP and IQ scores could be improved by increasing the language requirements of the DAP.
DRAWING TEST BACKGROUND

Intelligence

Intelligence is viewed as an overall, global concept (g) that is made up of several discrete, but additive abilities. What abilities combine to result in g is a question of great debate, and various theories have arisen in an attempt to define the small units that make up the larger concept of intelligence. One theory (Wechsler, 1974) holds that intellectual ability is made up of two separate abilities: verbal abilities and perceptual-motor, or spatial abilities. Wechsler believes that these two areas comprise the whole scope of one's abilities and that, when combined, reflect the individual's total, or global, intellectual ability level. Verbal abilities, according to Wechsler, are those skills necessary for speaking, communicating, and interacting with others. Additionally, verbal abilities include logical and critical thinking skills that allow the individual to problem-solve in relation to social situations. Perceptual motor or spatial abilities are those skills needed for problem-solving in relation to small parts combining to equal a whole, and what the relationship between the parts is. These skills are more visual organization skills. These two skill areas appear to touch on the individual's total range of knowledge and abilities.

Draw-a-Person

The drawing test most used for a quick screening of intelligence is the Goodenough-Harris Drawing Test (GH). Recently, a new version, the Draw-a-Person (DAP) (Naglieri, 1988), was introduced with updated norms. Although the old GH version has been widely used as a means of measuring intelligence, criticism has been directed at the scoring system used for the Draw-A-Woman subtest, the imprecision of the norms provided, ambiguity of the overall scoring criteria, and the lack of norms for the Self-Drawing subtest. Naglieri revised the test to address these complaints.
A new scoring system was devised by Naglieri to reduce the influence of current styles of dress and to make scoring as objective and efficient as possible. The DAP scoring system is organized into three major components: 1) criteria; 2) items; and 3) categories. There are 14 criteria including 12 parts of the body (arms, ears, eyes, feet, fingers, hair, head, legs, mouth, neck, nose, and trunk), placement of certain body parts (attachments), and clothing. The criteria are scored on a number of specific characteristics, or items (e.g., detail of eyes, presence of fingers, proportion of trunk, etc.) of the drawings. A total of 64 items are scored. The items are divided into four categories: 1) presence; 2) detail; 3) proportion; and 4) bonus (Naglieri, 1988). By providing detailed descriptions and examples of scoring, Naglieri succeeds in making the scoring of the DAP much more efficient and objective.

Separate and distinct raw scores for the Man, Woman, Self, and Total (standard score based on the sum of the Man, Woman and Self drawings) are calculated. These raw scores are then transformed into standard score equivalents with a mean of 100 and a standard deviation of 15.

Based on the 1980 census, the standardization sample (N = 2622) was broken down into the appropriate demographic characteristics for age, sex, geographic region, race, ethnicity, and socioeconomic status. Naglieri took the information gathered from the standardization sample and split the age groups into quarter-year intervals for ages 5 through 8 and half-year intervals for ages 9 through 10 so as to allow for the rapid developmental processes that occur during those six years, and then divided ages 11 through 17 into their appropriate one-year intervals. Altogether, 21 age categories are specified, with an n = 200 per each one-year level. Finally, Naglieri updated the normative data in 1984 by administering the DAP to the standardization sample, aged 5 to 17.
Naglieri's efforts to implement improved scoring procedures resulted in improved reliability over the GH. Coefficient alpha was calculated for each one-year group between the ages of 5 and 17 to find the DAP's internal consistency. The mean values for the Total score ranged from $\alpha = .83$ to $\alpha = .89$. The mean values for the Man, Woman, and Self scores were somewhat lower than the Total score, ranging from $\alpha = .56$ to $\alpha = .78$. Stability of the DAP scores was analyzed by retesting a subsample ($n = 112$, grades 1 through 7) of the standardization sample 4 weeks later. Because older children's DAP Total scores tend to be higher, due to more refined development, a correlation that included Total scores for all ages was not computed. Instead, correlations were computed by student grade in an attempt to minimize the effect of older children’s Total scores on an overall reliability coefficient. Test-retest correlations for the Total scores ranged from $\alpha = .60$ to $\alpha = .89$ (mean $\alpha = .75$). Again, the test-retest correlations for the individual tests of Man, Woman, and Self were generally lower with means of $\alpha = .70$, $\alpha = .65$, and $\alpha = .58$, respectively.

For interrater reliability, the new DAP scoring system was compared to the GH scoring system. Two examiners scored the drawings of two samples of elementary and junior high school children ($n = 89$) using the same scoring system. Each examiner, who was trained to use both the DAP and the GH, scored each drawing using both scoring systems. The product-moment reliabilities between the standard scores obtained by the two raters using the same scoring system were high for both the DAP and GH scoring systems: GH, $\alpha = .86-.94$; DAP, $\alpha = .86-.95$. This indicates high interrater reliability. Intrarater reliability was measured by having raters rescore the tests of the above study approximately 20 days later. Results indicated that both the new DAP and GH scoring systems are stable over time within raters: DAP, $\alpha = .89-.98$; GH, $\alpha = .93-.95$. 
To assess the new DAP's ability to actually measure what it purports to measure (construct validity), Naglieri used a criterion of developmental change in mean scores, under the premise that mean raw scores on the DAP should increase as the child's ability to produce human figures improves with age. Evidence of this age differentiation is provided by an obvious increase in the mean raw scores of the DAP drawings and in the means for the Total raw score—especially from age 5 through age 11.

Two independent, concurrent validity studies showed that the new DAP scoring system and the GH scoring system correlate highly. When one rater scored the drawings with both systems, a correlation of $r = .77-.82$ resulted. The second concurrent validity study involved scoring 100 drawings according to the two systems, but this time all drawings were scored by the same examiner using one scoring system, recording the scores on different protocols, scrambling the drawings, and then scoring according to the second system. The resulting correlations were higher than the first study ($r = .80-.87$). Naglieri (1988) asserts that these two studies indicate that the DAP and GH have a great deal of overlap in the construct they measure, but that the small differences in these correlations and the interrater reliabilities reported indicate a small difference in what each scoring system is measuring.

The DAP's relationship with a criterion variable (criterion validity) such as abstract figural analogies and achievement in reading and mathematics was also analyzed. When correlated with the Matrix Analogy Test-Shor Form (MAT-SF), "a nonverbal measure of ability which uses abstract figural analogies of the progressive matrix format," (Naglieri, 1988, p.18) the results indicated agreement of $r = .28-.31$ for grades K through 3, and $r = .19-.27$ for grades 4 through 12. However, the MAT-SF is a brand new test whose validity is still being evaluated, and whose definition of purpose is vague. A correlation between the DAP and the Multilevel Academic Survey Test (MAST), a
measure of school achievement, yielded correlations of $r = .19-.24$ for reading and $r = .17-.21$ for mathematics. Although Naglieri attempted to provide ample psychometric information, he falls short in the validity criteria by using the MAT-SF, an unsubstantiated measure of a poorly-defined concept named nonverbal ability. It is unclear what Naglieri means by nonverbal ability which throws the criterion validity information presented into question, but the more important construct validity is also suspect.

**Goodenough-Harris**

Although Naglieri provides quite a bit of his own research on the DAP, other independent research data is quite limited. Most research still pertains to the GH. Since the new DAP is based on the GH, however, GH information may provide some insight on the characteristics of the DAP. The Goodenough-Harris Draw-a-Man Test was originally published in 1926 by Florence Goodenough. Goodenough developed the 51 point-scale item scoring system to measure the developmental character of children's drawings. Prior to this instrument, the developmental character of children's drawings was discussed mainly via theoretical frameworks, but provided no avenue for measurement of the construct; Goodenough's instrument, in contrast, allowed for the quantified evaluation of children's actual freehand drawings. It was established by Goodenough that children's drawings involved an intellectual component that could be measured and related to the current psychometric studies of that time. Continued research showed that all children's drawings had a common element: the quality of the drawings improved in age increments in terms of coherence and the presence of more fully developed detail. Children's drawings were developmental in character.

In 1963, Dale Harris revised the Goodenough test and titled it the Goodenough-Harris Drawing Test (GH). The GH is defined as a measure of a child's
(age 5-14) current level of intellectual functioning based on his/her drawing of a man and a woman. Harris (1963a) proposed that the concept of intelligence be replaced with the idea of intellectual or conceptual maturity. He defined intellectual maturity as follows: 1) the ability to perceive; 2) the ability to classify objects into categories; and 3) the ability to generalize. These three abilities are said to make up the process of concept formation. As the child progresses in age, his/her quality of drawings improves. In other words, as the child ages chronologically, s/he exhibits signs of better concept formation, thus producing better human figure drawings.

Additionally, Harris replaced the outdated mental age/chronological age ratio IQ with the deviation IQ measure, lengthened the previous scoring system from 51 point-scale items to 73, and added the Draw-a-Woman Scale and the Self-Drawing Scale, which are scored in the same manner as the Draw-a-Man Test. [Although the Self-Drawing Scale was developed as a possible projective test, research has not supported this use.] Harris also added a set of Quality Scale cards was also added to the test packet to help users of the test more quickly score the child's drawings. The Quality Scale is made up of 24 cards of standard drawings to be compared with the child's drawings.

**PSYCHOMETRIC PROPERTIES/DIFFICULTIES**

Several studies have been conducted over the years to evaluate the GH's validity. Unfortunately, the results have been discouraging. Most studies have compared test scores on the GH to intelligence tests of well-known validity such as the WISC-R. When compared to WISC-R scores, the GH correlated with the Full Scale IQ score in the range of \( r = 0.63-0.73 \) (Tramill, Edwards, & Tramill, 1980; Dunn, 1967b; Naglieri & Maxwell, 1981). Correlations between the GH and WISC-R Verbal Scale IQ scores were in the range of \( r = 0.56-0.59 \) (Tramill, et al., 1980; Dunn, 1967b). WISC-R
Performance Scale IQ scores correlated in the range of $r = .55-.62$ with the GH (Tramill, et al., 1980; Dunn, 1967b; White, 1979). These correlations are all low, and although they are still statistically significant, the numbers fail to reach clinical significance when the variance is determined (Scott, 1981). A maximum of 53% of the variance was accounted for by the relation of the GH and the WISC-R Full Scale IQ, which are said to measure the same domain—intelligence. How well the Naglieri version, with improved reliability, correlates with a standard intelligence test, such as the WISC-R, is, as yet, unknown.

It has generally been assumed that the GH and the DAP are primarily tapping a child's general ability. The GH and the DAP, however, use a purely nonverbal method; that is, neither test requires vocal input from the subject. And although each test does, to an extent, tap verbal skills (as indicated by correlations with Verbal Scale IQ) increasing language demands could result in higher correlations with verbal and more popular tests of intelligence. Harris (1963b) proposed that one's intelligence, or intellectual/conceptual maturity, is made up of the abilities to perceive, classify, and generalize, which make up one's general ability level. These abilities can be tapped, to some degree, by the drawing test. What drawing tests fail to consider, however, is the role that language plays in intellectual abilities. Considerable research exists that demonstrates that there is a strong verbal component to intelligence as traditionally measured (Wechsler, 1974; Dunn, 1967a). Thus a more complete approach to measuring these intellectual abilities would be to add some language tasks, such as talking about the drawing. Language would allow the child to use classification and generalization skills in making their drawings. After all, a child's drawing never portrays an object exactly as it appears; the drawing is modified by the child (Frederickson 1985). The end product is a modified, added to, and selected drawing of an object that the child perceived. The
drawing is guided by the meaning of the object to the child. In this way language may be closely related to the child’s ability to draw. Indeed, one study (Rawl, 1968) that required subjects to verbalize while taking the GH resulted in improved GH scores. Because the GH requires only that the child draw, without talking about the drawing, it may fail to tap the language/verbal aspect of intelligence.

HYPOTHESES

Two questions thus arise after evaluating the theory and purposes behind drawing tests: First, what is the validity coefficient between the Naglieri DAP and an estimated Full Scale IQ score on the WISC-R? And, second, does adding verbalizations improve the validity coefficient of the DAP with an estimated Full Scale intelligence score? In an attempt to answer the first question, the DAP’s validity, it was correlated with an IQ estimate obtained from the Vocabulary subtest from the Verbal Scale of the WISC-R and the Block Design subtest from the Performance Scale of the WISC-R. These two subtests were chosen because they are the purest measures of their respective scales and constitute a commonly used short form. The full WISC-R test was not used due to the limited availability of subject time. Vocabulary correlates with the Verbal subscale the highest of all the verbal tests ($r = .86$), with the WISC-R’s Full Scale IQ Score at $r = .74$; the Block Design correlates with the Performance subscale the highest of all the Performance tests ($r = .68$), and with the Full Scale IQ Score at $r = .68$. This study first evaluates the validity of Naglieri’s DAP by correlating DAP Total scores with the estimated IQ scores obtained from the Vocabulary and Block Design subtests of the WISC-R. Correlations should be greater than earlier studies of the GH due to the improved reliability of the DAP. Studies that correlated GH scores and WISC-R Full Scale IQ scores yielded mean correlations of $r = .68$. A second hypothesis concerns the addition of student verbalizations while taking the DAP. It predicted that those subjects
who verbalize while taking the DAP would have higher correlations with their respective estimated Full Scale IQ scores than those subjects who did not verbalize while taking the DAP.
METHOD

Subjects

Subjects were selected for this study based on age. Only 9-, 10-, and 11-year-old males and females were allowed to participate since, as Naglieri (1988) indicates, this age range received the most consistent standard scores upon retesting. A total of 72 participants were tested in two private schools located in Bowling Green, KY (n = 41 from St. Joseph’s Catholic School; n = 4 from Anchored Christian School) and one private school located in Nashville, TN (n = 27 from St. Ann’s Catholic School). One hundred thirty-six permission slips were sent to all students in the desired age range. Only those students whose parents completed and returned the permission slip (Appendix A) participated in the study. Of the 72 students whose parents gave permission, 70 were actually given the research instruments. Two students were not tested due to their being ill during the times of testing. Eight of the 70 subjects who were tested were not considered when the data was compiled; their scores were not used due to their being undertested during the test administration. The final number of subjects was 62.

Of the subjects tested, 46% (n = 29) were male and 53% (n = 33) were female. The number of students in each age category were as follows: 31% (n = 19) were nine years old, 34% (n = 21) were 10 years old, and 35% (n = 22) were 11 years old. The demographic information obtained from the parents provided additional normative information. Eighty-five percent (n = 61) of the parents returning the demographic questionnaires (Appendix A) were females. Eighty-two percent (n = 59) of these females were married; 92% (n = 66) were Caucasian; 26% (n = 19) had a high school education, 32% (n = 23) an associates degree, and 25% (n = 16) a college education. These female
respondents had a mean age of 38 years and were primarily employed as managers and clerical workers (38%, n = 28).

Research Assistants

Four undergraduate students and one graduate student attending Western Kentucky University served as research assistants. The undergraduate students received course credit for their participation; all assistants were naive as to the purposes and expected results of the study. Two of the undergraduates and the graduate student served as scorers. The other two undergraduates served as examiners.

Research Assistant Training

For training purposes, all of the undergraduate students attended four instructional sessions, lasting 1 1/2 hours each, designed to teach either the administration or scoring procedures of the DAP, Vocabulary subtest and Block Design subtest, depending on the assistant’s role in the study. The graduate student had received previous training in administration and scoring of the instruments through graduate coursework. These sessions also provided the assistants with several practice sessions in which they were allowed to ask questions and practice prompts. The examiners were also trained concerning the cueing of the subjects in the Verbalization group. They were instructed to state each prompt at least one time during the administration. No further training was provided on this aspect and, as a result, the style of each examiner may become evident in the results of the study. Both examiners approached the cues in a different manner—one cued with both prompts at the onset of administration, and the other waited until the subject had been drawing for about 45 seconds before he issued the prompts. Since the subjects were assigned to the Verbalization group randomly, it was assumed that the examiner differences would equal out. All assistants, including the
graduate student, were required to pass a competency test prior to beginning the actual testing of subjects or scoring of protocols.

The competency test consisted of assistants scoring or administering one test packet. The packet consisted of one Vocabulary subtest, one Block Design subtest, and one DAP test. The competency test packet was previously scored by the author to identify the correct scores for each drawing or Vocabulary item score. To analyze scorer competency, each assistant’s total scores on the DAP, Vocabulary subtest, and Block Design subtest were correlated with the correct scores for the protocol. Prior to beginning actual subject testing, each assistant’s DAP, Vocabulary subtest, and Block Design subtest were required to correlate .95 with the criterion scores. If the correlation was not at .95, then additional instruction was provided to help the assistant obtain the cut-off correlation level and another competency test was administered. The .95 criterion correlation level was used in this study as it is a commonly used cut-off level for relationships that occur as a result of the effects of the independent variables rather than by chance.

The examiners were also given competency scores based on their performance. Each examiner tested a practice subject who followed on a script so s/he would respond in a consistent manner for each examiner. The examiner was observed by the project supervisor to determine if he began and ended the tests at proper points and questioned subject responses where appropriate. The Vocabulary subtest and Block Design subtest final scores were correlated with their criterion scores. If the examiner neglected to question vague responses, or stopped testing prior to the appropriate stopping point, his score would not correlate well with criterion scores. The DAP was not scored or correlated as the examiner only rehearsed the instructions and prompts with the practice subject. A correlation of .95 was also the competency level required for the examiners.
Materials

The Draw-a-Person (DAP) (Naglieri, 1988) and a short form of the Wechsler Intelligence Scale for Children-Revised (WISC-R) were used. The DAP is an untimed paper and pencil drawing test which requires the subject to draw a picture of a man, of a woman, and of the self.

The Vocabulary subtest from the Verbal Scale and the Block Design subtest from the Performance Scale from the WISC-R were used as the criterion measure. The Vocabulary subtest consists of 32 words arranged in order of increasing difficulty. The child is asked to orally explain the meaning of each word. The Vocabulary subtest is the best single subtest measure of g, or the Full Scale IQ score (64% of its variance may be attributed to g), is the most reliable subtest in the scale (r = .86), and has moderate correlations with the Full Scale (r = .74) and Verbal Scale (r = .68) (Wechsler, 1974). The Block Design subtest consists of 11 timed items. It requires the subject to use red and white blocks to assemble a design identical to a picture. The Block Design subtest is the best measure of g among the Performance Scale subtests and the fourth-best measure of g among all 10 subtests (53% of its variance may be attributed to g). Its reliability is good (r = .85) and it correlates moderately well with the Performance Scale (r = .68) and the Full Scale (r = .68) (Wechsler, 1974). The two scores on these subtests are then changed into scaled scores and are then added together. Sattler (1988) provides equivalent estimated IQ scores based on the summed scaled scores. These two subtests are popularly used as a short form of the WISC-R and were used because a reasonable estimate of intelligence can be obtained from the two, according to Sattler (1988). Also, administering only two subtests limited the subject’s lost time from classroom instruction. The two scores were then combined to form an estimated IQ score for the child by using Sattler’s (1988) formula.
**Procedure**

Examiners were provided with a list of participating subjects. Each subject was taken to the testing room by the examiner. The child was randomly assigned to either a Verbalization or Nonverbalization group according to the order of permission slips returned. Testing began when the child was taken from the classroom and brought to the testing room, which was designated by the school. Each examiner tested his subject individually. Both examiners followed the same procedure. Each examiner was provided with a set of testing packets that designated the order of subtests to be administered and the Verbalization/Nonverbalization assignment. This order was developed by assigning the DAP, Vocabulary, and Block Design subtests a number and then following the random number table to designate the order of presentation. The DAP Man, Woman, and Self subtests were also presented in a random order. The order of administration of the DAP and WISC-R (short form) was randomly assigned in each group to control for order effects. The Nonverbalization group followed standard administration procedures for the DAP, Vocabulary and Block Design subtests.

In the Verbal group administration, the DAP was administered under standard conditions, but with the examiner prompting the child with the statements, "Describe to me what you are doing," and "Tell me all about this picture." As the child discussed the drawing, the assistant tape-recorded the responses. If the child began to discuss topics other than the drawing, the assistant repeated the prompt statements in order to bring the child back on task. Administration of the Vocabulary and Block Design subtests followed standard procedure guidelines as outlined by Wechsler (1974). Each test was then scored by the assistants trained to score the instruments.
RESULTS

The validity coefficient for the traditional DAP Total score using estimated Full Scale IQ score as the criterion was low ($r = .33, p > .05$). Means and standard deviations for male and female DAP Total scores were calculated on the sample to determine if they were similar to the standardization population used in Naglieri's DAP standardization population. Comparison suggests that the sample population was similar to Naglieri's standardization sample. These data are presented in Table 1.

The second prediction concerned the addition of student verbalizations to the administration of the DAP. Based on the literature, an increase of verbal requirements on the DAP should result in improved correlations between the DAP Total Score and its corresponding estimated Full Scale IQ Score. This hypothesis was also not supported by the study's results. A Pearson Product-Moment correlation coefficient ($r$) was computed on the group of students taking the DAP under the Verbal condition; the result was an $r = .52, p > .05$ with estimated Full Scale IQ in comparison to the Nonverbal group's $r = .33, p > .05$ with estimated Full Scale IQ. The correlations between all tests for Verbal and Nonverbal groups are provided in Table 2. Although the verbalizations did increase the correlation, when a two-tailed test of difference between two independent correlation coefficients was performed, it yielded a nonsignificant difference ($t(60) = -.92, p > .05$). To test further the effects of the verbalizations on the DAP Total score, two-tailed t-test on the mean DAP Total scores (Verbal vs. Nonverbal) was computed. The result was a nonsignificant $t(60) = 1.12, p > .266$.

To determine if the effects of the amount of time the verbalization group talked had any relationship with the estimated Full Scale IQ, a Pearson Product-Moment correlation coefficient ($r$) was computed. It resulted in a nonsignificant $r = .15, p > .05$. 


Table 1

Means and Standard Deviations by Grade Level for Male and Female Draw-a-Person Total Scores

<table>
<thead>
<tr>
<th>Current Study</th>
<th>Standardization Sample*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sex</td>
</tr>
<tr>
<td></td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>11</td>
</tr>
</tbody>
</table>

* based on Naglieri's standardization sample
** n based on total standardization sample of 398
Table 2

Correlation Matrix for Verbal and Nonverbal Groups With All Tests

### Nonverbal Group

<table>
<thead>
<tr>
<th></th>
<th>VOC</th>
<th>BD</th>
<th>DAPT</th>
<th>DAPM</th>
<th>DAPW</th>
<th>DAPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSIQ</td>
<td>.755</td>
<td>.829</td>
<td>.326</td>
<td>.159</td>
<td>.325</td>
<td>.298</td>
</tr>
<tr>
<td>VOC</td>
<td>.261</td>
<td>.338</td>
<td>.158</td>
<td>.353</td>
<td>.171</td>
<td>.169</td>
</tr>
<tr>
<td>BD</td>
<td>.186</td>
<td>.088</td>
<td>.944</td>
<td>.757</td>
<td>.683</td>
<td>.772</td>
</tr>
<tr>
<td>DAPT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAPM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAPW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Verbal Group

<table>
<thead>
<tr>
<th></th>
<th>VOC</th>
<th>BD</th>
<th>DAPT</th>
<th>DAPM</th>
<th>DAPW</th>
<th>DAPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSIQ</td>
<td>.797</td>
<td>.860</td>
<td>.523</td>
<td>.409</td>
<td>.368</td>
<td>.588</td>
</tr>
<tr>
<td>VOC</td>
<td>.523</td>
<td>.445</td>
<td>.338</td>
<td>.227</td>
<td>.548</td>
<td>.457</td>
</tr>
<tr>
<td>BD</td>
<td>.394</td>
<td>.333</td>
<td>.286</td>
<td>.849</td>
<td>.723</td>
<td>.644</td>
</tr>
<tr>
<td>DAPT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAPM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAPW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FSIQ = estimated Full Scale IQ  
DAPM = DAP Man Score  
DAPW = DAP Woman Score  
DAPS = DAP Self Score  
DAPT = DAP Total Score  
VOC = Vocabulary Subtest  
BD = Block Design Subtest
Discussion

The first task of this study concerned the determination of a validity coefficient for Naglieri’s DAP as measured by an estimated Full Scale IQ score from the WISC-R (short form). Naglieri’s DAP lacked any validity information that related to widely-used and accepted intelligence measures, such as the WISC-R. To determine the validity coefficient, DAP Total scores were correlated with the estimated Full Scale IQ scores from the WISC-R (short form). It was predicted that the improved reliability of Naglieri’s DAP would result in a relationship between the two measures. The low correlation ($r = .33, p > .05$) is lower than coefficients found for GH and WISC-R Full Scale IQ scores. This suggests that although Naglieri’s current revision of the drawing test is purported to be more efficient and reliable, his restructuring has in some way altered the existing relationship between intelligence and the developmental characteristic of human figure drawings. Whereas the GH edition resulted in a mean correlation of .68 with WISC-R Full Scale IQ score, the correlation in this study dropped below the significance level.

Naglieri’s efforts to streamline the scoring system for efficiency and ease could also be a cause of the low correlation. In order to make scoring criteria objective and simple, Naglieri decreased the number of scorable items from 73 to 64. This revision also likely results in a decrease in the test’s reliability, and thus, validity, due to restriction of variability.

Another possible cause for this low correlation and, hence, the implied decrease in validity, is the instrument used to measure intelligence. The estimated Full Scale IQ score was obtained by administering the Vocabulary and Block Design subtests of the WISC-R. Although these two subtests are the strongest measures of their respective scales and correlate the best with the Full Scale IQ score, estimating intelligence based
on such a restrictive sampling of the individual's abilities results in error in the estimation of IQ which, in turn, decreases correlations. Based on only two samples of a rather complex and infinite range of responses available, the estimated IQ obtained from the abbreviated WISC-R is but a rough estimate of the global concept of intelligence. There are many abilities that combine to form intelligence, and this short form taps just two of the many. A full WISC-R would be a better measure of the ability level of the person tested.

Finally, the low correlation may have resulted from some other indefinite/undeterminable factor. It may be that the Naglieri drawing test is measuring some other unknown skill than what is assessed by the WISC-R.

Although the correlation between the Verbalization group's DAP Total score and the estimated Full Scale IQ score was not significant ($r = .52$), there was a trend toward an increase over the Nonverbal group ($r = .33$). Again, the small sample size may have prevented the results from being significant, seeing as this sample shared many of the same characteristics as the standardization population. With a larger sample size, it is probable that the correlations between the Verbal DAP and estimated Full Scale IQ scores would be significant. Still, this study's increase in the Verbal DAP's correlation with the estimated Full Scale IQ scores lends some credibility to the notion that it is the DAP's lack of a verbal component that restricts its correlations with WISC-R IQ scores.

Despite the interesting results of the study, several limitations were observed in this study and should be considered in carrying out future research. When undertaking a validity study for a proposed intelligence measure, the criterion instrument should be a validated instrument in and of itself. The more valid the criterion instrument, the more meaningful the predictor instrument's scores will be. For the purposes of this study, a short form of the WISC-R was used so as to provide the most time-efficient measure of
the DAP's ability to measure intelligence. Although this particular short form, consisting of the Vocabulary and Block Design subtests of the WISC-R, is a recommended short form to obtain an estimate of intelligence, it can provide only a very limited measure of ability. Sattler (1988) recommends that, whenever possible, the full intelligence scale be administered so as to provide the best measure possible. Even with a full WISC-R administered, the intelligence measurement will still only be an estimate of the person's full abilities. However, an estimate based on 10 subtests that all correlate well with the Full Scale IQ score is much closer to complete measurement than two subtests. Therefore, any future validity studies concerning the DAP's predictive ability of intelligence should implement a complete intelligence scale.

The sampling method used in the present study provided limited information as well, as it was restricted to private school children. Despite the accessibility of children attending private schools, they do not represent the population in general. Privately-schooled children's parents who pay for their child's education usually make more money, and their perspective on the educational process is different from that of parents who send their children to public schools. As a result, the privately-educated child is also different from the publicly-educated child. Any future studies on this topic should sample a broader range of children; it should also sample more than 60 children. A larger sample size is always desirable for any validity study, as it provides more information.

While the addition of verbalizations did show a trend toward better DAP-estimated Full Scale IQ score correlations, it was not a significant one. The examiner prompts used may be a source of trouble for the study's results. The prompts may need to be modified to encourage increased verbalizations from the subject. While the addition of verbalizations did show a trend toward better DAP-estimated Full Scale IQ correlations,
it was not a significant one. The examiners varied in their prompting of the children, despite the training. Upon listening to the tape recordings of examiner styles at the completion of data collection, it appeared as though the examiner who reissued the prompts more than one time elicited more verbal remarks from the subject. It is unclear, however, if this increased verbalization was due to ability on the part of the child or just a more relaxed test setting. By repeating the prompts at various periods during the administration, and not just at one particular time, the subjects were constantly forced to describe their actions, and, hence, concentrate on the task. Caution should continue to be used, however, as too frequent prompting from the examiner could be construed as leading the subject and would result in the subject being coerced into making up stories about the pictures to please the examiner. Increased verbal responses, then, may be less of a reflection of the subject's ability. This examiner difference is a good starting point for further research, as a study comparing the different styles of presentation could be useful in determining their effects on the DAP Verbalization group's correlations with the WISC-R (short form). The styles of the examiners should obviously be more closely monitored as the examiner difference discovered in this study puts limitations on the conclusions that can be drawn from the results. By not keeping examiner styles constant, the effects of added verbalizations are unclear. Also prompting at more consistent intervals, for example, at the beginning of testing, once after a minute has expired, and once after two minutes have gone by, may prove to be an interesting treatment effect.

Although this study provided no conclusive results concerning the effects of adding verbalizations to the DAP's intellectually predictive quality, it did provide useful information about the impact of verbal skills in predicting intelligence scores on the WISC-R. Although statistically insignificant, increasing verbal requirements did result in
an increase in DAP-WISC-R (short form) correlations. Research concerning this verbal aspect in drawing tests should continue, and should also explore other avenues for uncovering the role of verbal abilities in intelligence.

Despite disappointing correlations with a standard intelligence test, the human figure drawing test has frequently been used as a quick assessment of intelligence. Its popularity has been sustained over the years because of its relative ease of administration and rapid scoring potential. The DAP is the most recent edition of a human figure drawing test that began in 1926 by Goodenough. Naglieri’s edition also boasts of administration simplicity and quick, objective scoring. In this sense, Naglieri has maintained his human figure drawing test as an easy method of assessing children’s abilities. Despite Naglieri’s attempted improvements in the drawing test, he did not demonstrate adequate validity information by correlating his DAP with a standardized intelligence test. This study attempted to determine the validity coefficient for the DAP by using a short form of the WISC-R as a criterion measure. Based on the data from this study, Naglieri’s stated improvements in the drawing test failed to be reflected in the correlations with the WISC-R (short form). These data raise questions as to the validity of the DAP as a measure of intelligence, and naturally warrant further study with improved methods of data collection than demonstrated in the present study.

Perhaps the validity was unsubstantiated because Naglieri’s DAP keeps up the spatially-loaded orientation of the drawing test. The addition of a verbal requirement was evaluated in this study to determine its effects on correlations with intelligence scores. Although there appeared to be a trend toward improved correlations between the DAP and intelligence scores, it was not significant. If anything, this information raises further questions as to the DAP’s potential uses as a quick, valid measurement of the elusive concept called intelligence.
Permission Form

My son/daughter, ____________________________, has my permission to be tested by a psychology graduate student for the purposes of her research project. I understand that all results will be kept confidential and that my child’s name will not be associated with any results. Additionally, I am aware that I will be informed about the results of the project if I so desire.

Parent/Guardian __________________ Date ________________

Please circle the correct response or fill in the blanks.

1. Your Age______ Sex: Male       Female
2. Marital Status: single married divorced widowed
3. Spouse’s Age______
4. Race: Hispanic Black Caucasian Asian Other
5. Your Education: Years of School Completed___________
   Spouse’s Education: Years of School Completed_________
6. Your Occupation_____________________________________
   Spouse’s Occupation_________________________________

****If you desire project results to be mailed to you, please include your mailing address below.

________________________________________________________
Appendix B

WESTERN KENTUCKY UNIVERSITY
BOWLING GREEN, KENTUCKY 42101

Dear Parent,

I am a graduate student at Western Kentucky University and am conducting research on the methods of assessing children's intelligence. Through research, psychologists and school administrators can get a clearer understanding of a child's potential and abilities so that the best possible education becomes available. The goal of this project is to determine how adequate a particular test is in determining a child's intelligence. Our ability to achieve this goal depends on your willingness to complete the questionnaire and allow your child to participate in the study.

If you are willing, I would appreciate your taking five minutes to fill out the permission slip for your child and the background information questionnaire I have included and then returning them in the enclosed envelope. The background information is important in understanding how similar the people tested in this project are to the population in general. Your child may still participate if you choose not to complete this questionnaire. Yours and your children's privacy will be protected, as no names will be connected with test materials or results.

Please return the questionnaire and permission slip in the attached envelope and mail them to me within one week. Should you grant permission, your child will be tested during school. Because the child's name will not be attached to the results, I cannot give you specific information about how your child performed. I will be able to tell you the results of the research in general if you desire. If so, please include your mailing address so that the project results can be sent to you. Thank you for your time and effort in providing the background information and allowing your child to participate in this important research project.

Sincerely,

Phyllis H. Millsapugh
Psychology Graduate Student
Western Kentucky University
843-1199

Sally L. Kuhlenschmidt, Ph.D.
Thesis Supervisor
Assistant Professor
Western Kentucky University
745-4417
Subject #:  
Sex:  
Today's Date:  
Date of Birth:  
Age:  
Grade:  

VOCABULARY--discontinue after 5 consecutive failures

Score
2, 1, 0 pts.

__ 1. Knife_______________________________

__ 2. Umbrella____________________________

__ 3. Clock______________________________

8-10 yrs __ 4. Hat________________________

__ 5. Bicycle____________________________

11-13 yrs __ 6. Mail_______________________

__ 7. Alphabet____________________________

14-16 yrs __ 8. Donkey_______________________

__ 9. Thief______________________________

__ 10. Join_______________________________

__ 11. Brave______________________________
12. Diamond

13. Gamble

14. Nonsense

15. Prevent

16. Contagious

17. Nuisance

18. Fable

19. Hazardous

20. Migrate

21. Stanza

22. Seclude

23. Mantis

24. Espionage
___ 25. Belfry

___ 26. Rivalry

___ 27. Amendment

___ 28. Compel

___ 29. Affliction

___ 30. Obliterate

___ 31. Imminent

___ 32. Dilatory

Total Score: _______
## APPENDIX D

**Subject #:**
**Sex:**
**Today’s Date:**
**Date of Birth:**
**Age:**
**Grade:**

### BLOCK DESIGN TEST—Discontinue after 2 consecutive failures

<table>
<thead>
<tr>
<th>Design</th>
<th>Time</th>
<th>Pass-Fail</th>
<th>Score (Circle the appropriate score for each design)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Design 1" /> 1. 45”</td>
<td><img src="image2" alt="Time 1" /> 1 2</td>
<td><img src="image3" alt="Pass-Fail 1" /> 0 1</td>
<td><img src="image4" alt="Score 1" /> 2</td>
</tr>
<tr>
<td><img src="image1" alt="Design 1" /> 2. 45”</td>
<td><img src="image2" alt="Time 1" /> 1 2</td>
<td><img src="image3" alt="Pass-Fail 1" /> 0 1</td>
<td><img src="image4" alt="Score 1" /> 2</td>
</tr>
<tr>
<td><img src="image1" alt="Design 1" /> 8-16 yrs. 3. 45”</td>
<td><img src="image2" alt="Time 1" /> 1 2</td>
<td><img src="image3" alt="Pass-Fail 1" /> 0 1</td>
<td><img src="image4" alt="Score 1" /> 2</td>
</tr>
<tr>
<td>4. 45”</td>
<td></td>
<td></td>
<td><img src="image4" alt="Score 1" /> 0 4 5 6 7</td>
</tr>
<tr>
<td>5. 75”</td>
<td></td>
<td></td>
<td><img src="image4" alt="Score 1" /> 0 4 5 6 7</td>
</tr>
<tr>
<td>6. 75”</td>
<td></td>
<td></td>
<td><img src="image4" alt="Score 1" /> 0 4 5 6 7</td>
</tr>
<tr>
<td>7. 75”</td>
<td></td>
<td></td>
<td><img src="image4" alt="Score 1" /> 0 4 5 6 7</td>
</tr>
<tr>
<td>8. 75”</td>
<td></td>
<td></td>
<td><img src="image4" alt="Score 1" /> 0 4 5 6 7</td>
</tr>
<tr>
<td>9. 120”</td>
<td></td>
<td></td>
<td><img src="image4" alt="Score 1" /> 0 4 5 6 7</td>
</tr>
<tr>
<td>10. 120”</td>
<td></td>
<td></td>
<td><img src="image4" alt="Score 1" /> 0 4 5 6 7</td>
</tr>
<tr>
<td>11. 120”</td>
<td></td>
<td></td>
<td><img src="image4" alt="Score 1" /> 0 4 5 6 7</td>
</tr>
</tbody>
</table>

**Total:**
DAP WORKSHEET--MAN

Directions
Say: I’d like you to draw some pictures for me. First, I’d like you to draw a picture of a man. Make the very best picture you can. Take your time and work very carefully, and I’ll tell you when to stop. Remember, be sure to draw the whole man. Please begin. (allow 5 min.)

1. ARMS
   a) Presence   b) Detail 1   c) Detail 2
   d) Proportion   e) Bonus
   Total: ___

2. ATTACHMENT
   a) Attach 1   b) Attach 2   c) Attach 3
   d) Attach 4   e) Bonus
   Total: ___

3. CLOTHING
   a) Presence   b) Detail 1   c) Detail 2
   d) Detail 3   e) Bonus
   Total: ___

4. EARS
   a) Presence   b) Detail 1   c) Detail 2
   d) Proportion   e) Bonus
   Total: ___

5. EYES
   a) Presence   b) Detail 1   c) Detail 2
   d) Proportion   e) Bonus
   Total: ___

6. FEET
   a) Presence   b) Detail 1   c) Detail 2
   d) Proportion   e) Bonus
   Total: ___

7. FINGERS
   a) Presence   b) Detail 1   c) Detail 2   d) Detail 3
   e) Proportion 1   f) Proportion 2   g) Bonus
   Total: ___

8. HAIR
   a) Presence   b) Detail 1   c) Detail 2
   d) Bonus
   Total: ___
9. HEAD
a) Presence _____ b) Proportion _____ c) Bonus _____

10. LEGS
a) Presence _____ b) Detail _____ c) Proportion _____
d) Bonus _____

11. MOUTH
a) Presence _____ b) Detail _____ c) Proportion _____
d) Bonus _____

12. NECK
a) Presence _____ b) Detail 1 _____ c) Detail 2 _____
d) Bonus _____

13. NOSE
a) Presence _____ b) Detail _____ c) Proportion _____
d) Bonus _____

14. TRUNK
a) Presence _____ b) Detail _____ c) Proportion _____
d) Bonus _____

WORKING TIME: __________

TOTAL MAN
RAW SCORE: ________
Appendix F

Subject #: ______ Sex: ______
Today’s Date: ____________
Date of Birth: ____________
Age: ____________
Grade: ____________

DAP WORKSHEET--WOMAN

Directions
Say: This time I want you to draw a picture of a woman. Make the very best picture you can. Take your time and work very carefully, and I’ll tell you when to stop. Be sure to draw the whole woman. Please begin. (allow 5 min.)

1. ARMS
   a) Presence ___ b) Detail 1 ___ c) Detail 2 ___
   d) Proportion ___ e) Bonus ___
   Total: ___

2. ATTACHMENT
   a) Attach 1 ___ b) Attach 2 ___ c) Attach 3 ___
   d) Attach 4 ___ e) Bonus ___
   Total: ___

3. CLOTHING
   a) Presence ___ b) Detail 1 ___ c) Detail 2 ___
   d) Detail 3 ___ e) Bonus ___
   Total: ___

4. EARS
   a) Presence ___ b) Detail 1 ___ c) Detail 2 ___
   d) Proportion ___ e) Bonus ___
   Total: ___

5. EYES
   a) Presence ___ b) Detail 1 ___ c) Detail 2 ___
   d) Proportion ___ e) Bonus ___
   Total: ___

6. FEET
   a) Presence ___ b) Detail 1 ___ c) Detail 2 ___
   d) Proportion ___ e) Bonus ___
   Total: ___

7. FINGERS
   a) Presence ___ b) Detail 1 ___ c) Detail 2 ___ d) Detail 3 ___
   e) Proportion 1 ___ f) Proportion 2 ___ g) Bonus ___
   Total: ___

8. HAIR
   a) Presence ___ b) Detail 1 ___ c) Detail 2 ___
   d) Bonus ___
   Total: ___
9. **HEAD**
   a) Presence ___  b) Proportion ___  c) Bonus ___  
   Total: ___

10. **LEGS**
    a) Presence ___  b) Detail ___  c) Proportion ___  
    d) Bonus ___  
    Total: ___

11. **MOUTH**
    a) Presence ___  b) Detail ___  c) Proportion ___  
    d) Bonus ___  
    Total: ___

12. **NECK**
    a) Presence ___  b) Detail 1 ___  c) Detail 2 ___  
    d) Bonus ___  
    Total: ___

13. **NOSE**
    a) Presence ___  b) Detail ___  c) Proportion ___  
    d) Bonus ___  
    Total: ___

14. **TRUNK**
    a) Presence ___  b) Detail ___  c) Proportion ___  
    d) Bonus ___  
    Total: ___

WORKING TIME: ________

TOTAL WOMAN
RAW SCORE _____
DAP WORKSHEET--SELF

Directions
Say: Now I'd like you to draw a picture of yourself. Be sure to draw the very best picture you can. Take your time and work very carefully, and I'll tell you when to stop. Be sure to draw your whole self. Please begin. (allow 5 min.)

1. ARMS
a) Presence ___ b) Detail 1 ___ c) Detail 2 ___
d) Proportion ___ e) Bonus ___

2. ATTACHMENT
a) Attach 1 ___ b) Attach 2 ___ c) Attach 3 ___
d) Attach 4 ___ e) Bonus ___

3. CLOTHING
a) Presence ___ b) Detail 1 ___ c) Detail 2 ___
d) Detail 3 ___ e) Bonus ___

4. EARS
a) Presence ___ b) Detail 1 ___ c) Detail 2 ___
d) Proportion ___ e) Bonus ___

5. EYES
a) Presence ___ b) Detail 1 ___ c) Detail 2 ___
d) Proportion ___ e) Bonus ___

6. FEET
a) Presence ___ b) Detail 1 ___ c) Detail 2 ___
d) Proportion ___ e) Bonus ___

7. FINGERS
a) Presence ___ b) Detail 1 ___ c) Detail 2 ___
d) Detail 3 ___ e) Proportion 1 ___ f) Proportion 2 ___ g) Bonus ___

8. HAIR
a) Presence ___ b) Detail 1 ___ c) Detail 2 ___
d) Bonus ___
9. **HEAD**  
a) Presence  
b) Proportion  
c) Bonus  

Total:

10. **LEGS**  
a) Presence  
b) Detail  
c) Proportion  

Total:

d) Bonus

11. **MOUTH**  
a) Presence  
b) Detail  
c) Proportion  

Total:

d) Bonus

12. **NECK**  
a) Presence  
b) Detail 1  
c) Detail 2  

Total:

d) Bonus

13. **NOSE**  
a) Presence  
b) Detail  
c) Proportion  

Total:

d) Bonus

14. **TRUNK**  
a) Presence  
b) Detail  
c) Proportion  

Total:

d) Bonus

**WORKING TIME:**

**TOTAL SELF RAW SCORE:**
REFERENCES
References


