1 OF 1

ED

032932

DOCUMENT RESUME

ED 032 932 PS 002 157

By Meyer, William J., and Others
Measuring Perceptual Motor Ability in Preschool Children.
National Lab. on Early Childhood.
Spons. Agency-Office of Education (DHEW), Washington, D. C. Div. of Educational Labs.
Report No. - NLECE-70706-S-A-O-Y-10
Pub. Date (69)
Contract - OEC-3-7-70706-3118
Note - 17p.
EDRS Price MF - $0.25 HC - $0.95

Descriptors - *Cognitive Ability, Eye Hand Coordination, Factor Analysis, Factor Structure, Gifted, Group Tests, *Measurement Techniques, *Perceptual Motor Coordination, Predictive Ability (Testing), *Preschool Children, Test Interpretations, Test Reliability

Identifiers - Draw a Circle Task, Draw a Line Task, Preschool Abilities Test, Stanford Binet

A general working model of cognitive development assumes that there are sets of orthogonal cognitive abilities, which remain fairly stable after age 7. This paper examines the long term predictive and diagnostic value of assessing specific cognitive abilities among preschool children. This model by empirical studies was defendable on the grounds that the methodology of group empirical studies tended to prejudice results in favor of a general cognitive ability model. The assessment techniques used in this study draw heavily from a perceptual survey rating scale developed by Kephart for primary grades. Tests were administered to 74 middle class nursery school, 4- and 5-year-olds. The tests consisted of three visual pursuit tasks; measures of convergence, refixation ability, and power; and power; and the Draw-A-Circle task. In addition, the preschoolers were administered the Stanford-Binet, a specially developed preschool achievement test, and a measure of impulsivity control. Data was factor analyzed. Several problems identified were lack of observer agreement, unclearness as to what several of the tests were actually measuring, and scoring difficulties on the Draw-A-Circle. (MH)

Document Number 70706-S-A-0-Y-10

Measuring Perceptual Motor Ability in Preschool Children

William J. Meyer, Harvey Cohen,
David Goldstein

(Syracuse)

The research or work reported herein was performed pursuant to a contract with the Office of Education, U. S. Department of Health, Education, and Welfare through the Syracuse Center for Research and Development in Early Childhood Education, a component of the National Laboratory on Early Childhood Education, contract OEC-3-7-70706-3118.

MEASURING PERCEPTUAL MOTOR ABILITY IN PRESCHOOL CHILDREN

William J. Meyer, Harvey Cohen, David Goldstein

This paper is only peripherally concerned with the further development of talented children. The research program which will be described, along with some data, is however related to the identification of gifted children as well as the identification of young children with potential disabilities.

As a point of departure, it might be well for me to describe my general working model of cognitive development and behavior. Admittedly, this model is not especially unique to me nor do I believe that the model contributes anything really new to already existing sets of constructs. Perhaps, the unique contribution can be found in the level of analysis required by the model which I will attempt to show will substantially further our understanding of cognitive development and behavior.

A fundamental assumption of the model is that there exists sets of orthogonal cognitive abilities. It is further assumed that the nature of these orthogonal abilities remain relatively constant after seven years of age. I warned you that this model was hardly unique, but you are probably wondering why I am clinging to a formulation which has been somewhat badly mauled in terms of empirical analyses. My response is that these studies have relied almost entirely on group analyses, group testing procedures, and, most importantly, a unit of measurement relying entirely on the appropriateness of the response as opposed to the process by which the response was evoked. I shall now examine each of these points.

The Group as the Unit

You are all familiar with the standard procedure in studies of this sort; namely, a large group of children is administered varying numbers of tests comprising the battery of interest to the investigator. A matrix of intercorrelations is determined and from this is derived a factor structure. The investigator may, for theoretical reasons, decide a priori to achieve an orthogonal solution to the factor analysis. He is, of course, immediately open to the criticism, or at least the possible criticism, that an oblique solution was equally meaningful and that in all probability the correlation among factors would have been substantial. I frankly have little quarrel with this position because I essentially accept the notion of generalized intellectual ability or "g" when one is asking a question about general population. But it should also be noted that in group statistics it is very simple to overlook individual patterns of cognitive abilities. Stated somewhat differently, it is conceivable that a significant percentage of individual children do not conform perfectly to the "g" model but, rather, are seriously deficient in one or another of the several areas of cognitive capability. In essence, my argument is that the unit of analysis is inappropriate and serves to mask meaningful between subjects variations in cognitive patterns.

Group Administration of Measures. This variable is so obvious as to warrant only a brief comment. Clearly, when tests are administered in groups, and where multiple tests are administered within the same testing period, there is a significant source of common variation over measures which increases the probability of finding significant interability correlations. The work of Meyers, et al, strongly suggests that the preferential procedure is to administer the test battery on an individual basis using multiple examiners and counterbalancing testing sequences.

The Behavioral Unit. The most obvious and the simplest strategy in collecting data for factor analytic studies is the administration of measures which yield a relatively simple, straightforward numerical score. Certainly one cannot quarrel with the use of scores which presumably index a specific behavioral attribute. The question is what attributes are being indexed? Thus on any one task a child may respond before he has been able to perceive all of the attributes of the problem. He makes errors over a broad variety of tasks which is then interpreted as a lack of general intellectual ability but which in fact is attributable to a cognitive style. Another child may fail tasks which are timed, not necessarily because he is unable to perform the tasks but because he fails to achieve the time criterion. To conclude that this child does not possess the capability required by the task is misleading. Finally, in my own work with children between four and eight years of age, I am becoming convinced that the labels we use for our measurements, such as picture vocabulary, visual discrimination, and so forth, are little understood; that is, we know very little about the meaning of such behaviors in relation to levels of cognitive functioning as described by Piaget.

Returning now to the issue of identifying gifted children, I think it is perhaps safe to assert that the truly gifted child tends to make our assessment procedures look good. Thus he performs well on the tasks we give him and very often our assumptions about his performance are accurate. But as one surveys the literature on the gifted there is some cause for anxiety when it is discovered that there are studies on "gifted underachievers." Curiously, these studies implicitly assume that a child with a high general I.Q. functions equally well in all cognitive abilities. Thus the search is on for non-cognitive variables. I would propose that some effort be given to a finer analysis of the individual's capabilities and I would further assert that in order to do this meaningfully the research strategy should rely on a longitudinal approach. This seems crucial because the child may have experienced a cognitive or learning disability earlier in his life which developmentally corrects itself but which nevertheless has its residual effects on behavior at a later stage. This can occur when the child develops compensatory mechanisms or strategies for dealing with cognitive demands but where the strategies are no longer required. This study I am about to describe to you now is in its second year and we will not have additional follow-up data on the children until the end of May of this year. The intent of the study is to examine the long-term predictive and diagnostic value of assessing specific cognitive abilities among preschool children.

Kephart developed a perceptual survey rating scale designed to assess the perceptual-motor abilities of slow learners. Norms and scoring criteria were developed for primary school children but to my knowledge no work has been done in developing similar procedures for preschool children. This seems particularly unfortunate if there is concern about preventing failures and preventing children from developing inappropriate compensatory behaviors. It should also be noted that in the identification of gifted children, reliance on purely verbal measures may present a somewhat distorted picture of the child's capability.

We did not use all of the measures described by Kephart but rather selected those which, for one reason or another, seemed most promising for our purposes. Tasks requiring complex verbal instructions were avoided and tasks which seemed overly complex, thus reducing the variance, were also not included. In the final analysis, however, the selected tasks represent arbitrary choices. The tasks included then were: (1) three visual pursuit tasks involving left eye, right eye, and both eyes; (2) a measure of convergence; (3) refixation ability; (4) power; (5) draw-a-circle with preferred hand; and (6) draw-a-circle with both hands. In addition, the children were given the Stanford-Binet, a specially developed preschool achievement test, and a measure of the ability to control motor impulses.

Subjects. The original sample consisted of 92 children from middle class families, attending a university supported laboratory nursery school. A total of 18 children were not available for all of the tests administered and were, therefore, not included in the data analysis. The final sample is comprised of 32 males and 42 females for a total N of 74. The mean chronological age of the children was 56.32 months, with a standard deviation of 4.23. The mean Stanford-Binet (Form LM) I.Q. was 109.65 with a standard deviation of 12.73.

Procedure. Of the six oculomotor (OM) tests used, pursuits (left eye, right eye, and both eyes) and convergence as well as the 1-4 scoring system were adopted from Rosch and Kephart (1966), with the following modifications: (1) in the three pursuit tasks, otherwise adequate pursuit movements which proceeded in a series of small, distinct refixations were specifically defined as having a score of "2"; (2) for all visual tasks, S faced E across of a 2-1/2' high table. Two sponge pads were fastened to the surface of the table so that S could rest his elbows on the pads and rest his head on his hands. This effectively kept the head upright, facing forward and still; (3) two additional OM tasks were used; power-starting 20" from S's eyes, E moves penlight horizontally to bridge of S's nose. Score is the approximate number of inches between target and bridge of nose when one of S's eyes breaks from target; refixation--E places his hands at S's eye level, approximately 20" in front of S, and 45 deg. to either side of S. He then says: "I'm going to snap my fingers on one hand, then on the other, like this (demonstrates). Look where you hear the snap, but don't move your head." Hands are snapped alternately at 1/2 second to 1 second intervals. A score of "4" reflects ability to refixate smoothly, in one motion, and with no head movement. "3" indicates smooth movement but with one intermediate refixation. "2" reflects more than one pause or a long pause, or movement from one hand to the area of the second with a quick 'search' before actual refixation. "1" indicates inability to refixate without prolonged search.

The chalkboard tasks, draw-a-circle with one hand (DAC-1) and draw-a-circle with both hands (DAC-2) were also adopted from Roach and Kephart. The DAC-1 task required the child to simply take a piece of chalk, approach the chalkboard, and draw a circle. A score of "4" was recorded if the child's performance met the following criteria: proper shape, appropriate directionality (counter-clockwise for right-handed Ss), proper size 20-24 inches, and a crossing of the child's midline. A circle nearly perfect in size and shape received a score of "3." If either the midline was not crossed or the directionality was inappropriate, the child received a score of "2." If none of the criteria were met, a score of "1" was given to the child. The DAC-2 task required the child to take a piece of chalk in each hand, look directly at an "x" placed at his eye level, and draw circles with each hand. Unlike the DAC-1 task, the child was expected to draw several synchronous circles. A score of "4" was given if performance was smooth. If performance was adequate but stiff, a score of "3" was given. Extreme difficulty or inappropriate directionality resulted in a score of "2," while an inability to perform at all was given a score of "1." Additional modifications to the Kephart procedure were:

(1) The distance between S and the chalkboard was determined by having S walk directly forward with his elbow extended until the elbow touched the board.

(2) For the DAC-1 task, E held out the chalk directly in front of S so as not to bias the S's hand preference.

(3) The size of the circle deemed acceptable for the DAC-1 task was reduced for the preschool Ss (15-18 inches).

The Stanford-Binet, the Pre-School Abilities (PAT) Test, and the Draw-a-Line (DAL) tasks were routinely administered to all Ss at the laboratory nursery. The mean CA of the males was 50.94 with an SD of 12.08, while the mean CA of the females was 52.29 with an SD of 10.15. The mean I.Q. for males was 105.72 with an SD of 11.18; the mean I.Q. for females was 112.64 with an SD of 13.14.

The PAT is a device created to evaluate the curriculum of preschool programs and is not intended as an assessment device for individual preschool children. Items on the PAT seek to assess cognitive and perceptual motor abilities. We have not yet developed reliability estimates. The mean PAT score for males was 50.94 with an SD of 12.08, while the mean PAT score for females was 52.28 with an SD of 10.15.

The DAL is an index of impulsivity, based on Massari, Hayweiser, and Meyer (1969). A sheet of paper with an "X" placed at the top and bottom is used. The child is instructed to draw a line as slowly as he can between the two marks. A rate measure is obtained by dividing the length of the line by the time to draw the line. The inter-trial reliability after three administrations was .89. The mean DAL score for males was 1.02 with an SD of 1.30. The mean DAL score for females was 0.83 with an SD of 0.70.

Results

One of the purposes of this study was to develop procedures which would permit good observer agreement. This proved to be a formidable task. Indeed, it was our original plan to include the Walk-a-Beam task but this proved impossible because of the problem of observer agreement. I might add parenthetically that the problem has now been resolved, and we are including the task in our replication study currently underway.

The interjudge reliability for the power task was the highest of all the oculomotor assessments, the product movement $\underline{r}$ was .76; convergence produced the lowest correlation, $\underline{r}$ = .32 (see Table 1 for additional reliability data). Interjudge reliabilities for the DAC-1 and DAC-2 tasks were .92 and .67, respectively. Although these correlations are not, in general, impressively high, they do suggest that there are behavioral indicators involved in the tasks which are specifiable and observable. We have not yet been able to assess behavioral stability, but will have data available this coming summer.

A Varimax factor analysis was performed using the oculomotor measures, chalkboard measures, I.Q., sex, age, impulsivity (motor control), Preschool Achievement Test, and hand preference at the chalkboard. The resulting 6 factors accounted for 73% of the variance. Table 2 summarizes the factor loadings.

Discussion

The reliability coefficients for the OM measures are based on a sample taken during the training of the actual examiner, and should be viewed as conservative. Simply reading the scoring criteria was not adequate for proper scoring, and four or five hours of supervised practice were required.

However, in an unpublished pilot study (Cohen, 1966), employing a similar set of OM measures and one examiner who had simply read the scoring criteria, the overall interjudge reliability was .85. In the former case the examiner had no previous experience in the critical observation of visual behavior; in the latter case the examiner, although unfamiliar with the particular instrument, was studying special education for the visually handicapped. It is probably best, then, for an examiner to have some background experience and some practice with the instrument.

Factor I indicates the superiority of the females in a gross-motor task using the preferred hand, as well as in ocular pursuits. Factor IV, though, shows that younger females do worse on the same gross-motor task. Taken together, they indicate that, for the females, the sample covers a critical period in gross-motor development that some (Factor I) have passed while others (Factor IV) have not. The fact that younger females were noticeably inferior to males in the gross-motor task but not in ocular pursuits may be a reflection of the greater cultural emphasis on gross-motor skills and experience for males. In ocular pursuits females are apparently superior to males within the age range of the sample.

Factor V, involving target-to-target refixation and control of impulsivity, may be interpreted in three ways. (1) It is possible that the testing situation excessively penalizes the impulsive S. It is certainly reasonable to suppose that the highly impulsive child will be distracted and interrupted in the process of refixation. Further investigation is needed, in which eye movements from one target to another across a completely blank field are observed. (2) The refixation measure might tap no particular skill and therefore measure only the subject's willingness to tolerate the testing situation. The test still would be (and is) reliable, but it would not be valid. (3) The type of impulsivity assessed by the DAL could be a behavioral adaptation to poor refixation skills. That is, a child who is unable to refixate adequately is unable to map the space around him, so as to find a known target without a complete search. Once his fixation upon a task is interrupted or distracted, he is effectively "lost," and return to the task is difficult and effortful. Such a child might develop a response style involving completion of any task as fast as possible, before any interruption can occur. If this last interpretation is at least partially correct, it is possible that perceptual training designed to help the child learn to map the space around him would, if begun sufficiently early, preclude this type of response set.

The second interpretation of Factor V, that the refixation task reflects nothing but test-taking set, is of considerable methodological interest whether it is true or not. It would be profitable for many studies involving new testing applications to take the small amount of time involving new testing applications to take the small amount of time necessary to administer the DAL. If, in subsequent factor analysis, some measure(s) loads with DAL and nothing else, the investigator should consider the possibility that the measure--no matter how reliable--reflects nothing except the subject's willingness to tolerate the testing situation.

Factor VI suggests that the convergence test taps little more than the muscular ability to converge the eyes on a near target. Within the range of the sample, this is apparently largely irrelevant to P-M development, or to the classroom skill characterized by Factor II.

Factor III involves DAC-2 and "handedness," which was scored according to the hand-preference of S on DAC-1. Performance on DAC-2 that otherwise would have been given the maximum rating of "4" was penalized two points if the circles were drawn in the "wrong" direction for a person with that hand-preference. If a child happened to draw the single circle with his left hand, he was not penalized on DAC-2 for drawing the right circle clockwise and the left counter-clockwise, while the child who did DAC-1 with his right hand was. Since these 'left-handed' Ss did exceptionally well (i.e., were penalized less often), right-circle-clockwise/left-counterclockwise performance is probably at least as typical of the developmental level of the sample as is right-counterclockwise/left-clockwise. In any case, the indefinite hand-preference of preschool Ss makes it advisable to severely reduce or eliminate the penalty for "incorrect directionality" on DAC-2. If this had been done for the data under consideration, Factor III would probably have dropped out, and DAC-2 loaded on the same factors as DAC-1.

Factor II reflects test or classroom oriented skill. In view of the similarity between the PAT and the Binet, it is hardly surprising that they load together. It is important, however, that the highest loading on that factor is DAL. The implication is that impulsivity strongly influences test or classroom ability. One of the two conclusions follows from this. On the one hand, the DAL may be assessing a kind of "cognitive style," in which case educators may well have to work around the child's impulsivity. On the other hand, the DAL could be indexing a learned adaptive behavior (a possibility considered in the discussion of Factor V), in which case the component of classroom performance affected by impulsivity might be improved through early, appropriate perceptual training. Factor II thus underscores the importance of further investigation of the meaning and antecedents of the DAL score.

Table 1

Mean, Standard Deviation, and Interjudge Reliability for OM and DAC Tasks

Task	Mean	Standard Deviation	Interjudge Reliability
OM 1 (both eyes)	2.43	.97	.77
OM 2 (left eye)	3.08	1.13	.78
OM 3 (right eye)	2.59	1.01	.89
Convergence	3.18	.88	.32
Power	.86*	.96	.76
Refixation	3.58	.55	.48
DAC 1	2.20	.84	.92
DAC 2	2.00	.70	.67

* measured in inches

Table 2

Factor I		Factor II		Factor III	
OM Pursuits (1, 2, 3)	+	* DAL	+	DAC 2	+
Female	+	PAT	-	Left handedness	+
DAC 1	+	I.Q.	-		
24.8%		13.2%		9.4%	

Factor IV		Factor V		Factor VI	
Female	+	Refixation	+	Convergence	-
Age	-	DAL	-	**Power	+
8.8%		7.8%		7.6%	

* High score reflects inability to inhibit movement
** High score reflects inability to keep eyes converged on a near target