DOCUMENT RESUME

ED 135 157 EC 092 985

AUTHOR Mardell, Carol; Goldenberg, Dorothea S.
TITLE The Predictive Validation of a Pre-Kindergarten Screening Test.
PUB DATE Aug 76
NOTE 21p.; Paper presented at the International Scientific Conference of IFLD (3rd, Montreal, Canada, August 9-13, 1976) ; Best avaialble copy

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Educational Research; Identification; *Learning Disabilities; *Predictive Validity; *Preschool Education; Preschool Tests; *Screening Tests; Test Reliability; Test Validity
IDENTIFIERS Developmental Indicators Assessment Learning

ABSTRACT

Described is the development and revision over a 3-year period of the Developmental Indicators for the Assessment of Learning (DIAL), a screening test to identify pre-kindergarten children with learning disabilities. The DIAL is said to provide for gross motor, fine motor, cognitive, and communications assessment of large groups of children. Among the three years of DIAL research efforts discussed are a longitudinal study of 520 children, establishment of male and female norms for a population of 3100 children, and comparison of achievement and readiness scores with DIAL scores for 249 children. Among conclusions cited are high interrater reliability on the DIAL, significant correlation between DIAL scores and achievement and readiness test scores, and the feasibility of identifying pre-kindergarten children who need further evaluation. (CL)

THE PREDICTIVE VALIDATION of a PRE-KINDERGARTEN SCREENING TEST

Carol Mardell, Ph.D.

Dorothea S. Goldenberg, M.A.

BEST COPY AVAILABLE

There is a trend in education to identify children at younger and younger ages who may have trouble learning once they arrive at school. The question becomes deciding which tasks will correctly indicate the children who may be helped, thus avoiding future failure.

One way to select children with potential learning problems is to screen them, that is to use a test that will separate those children appearing to be developing in a satisfactory manner from those children experiencing developmental delay. Such a test unit must be one that can be routinely administered by an examiner with special limited training to large numbers of children in a relatively short period of time at a modest cost. In addition, a screening test cannot stand alone and may not be used for purposes of placement, intervention or treatment. Its purpose is to identify those children who need a diagnostic evaluation.

In attempting to provide a test for selection of pre-kindergarten children with potential learning disabilities, the DIAL battery was assembled. Its theoretical base lies in an assembly of developmental skills which encompasses all domains of learning.

Extensive research and experience with preschool and school age children led to the compilation of tasks and items for st . The Developmental Indicators for the Assessment of Learning, (DIAL), (Mardell and Goldenberg, 1972a) was written to meet the following criteria:

1. It must screen rather than diagnose.
2. It must be appropriate for 2½ - 5½ year old children.
3. It must be administered on a one-to-one basis.
4. It must be short in duration and paced to hold a young child's attention.
5. It must be multidimensional.
6. It must be noncategorical.
7. It must be scored objectively.
8. It must be process as well as product oriented.
9. It must be applicable to culturally different groups of children.
10. It must be normed on a stratified sample.
11. It must be economically feasible.

The outcome was an evaluation taking 25 - 30 minutes whereby a child moves through successive stations of DIAL which allow an observable recording of performance. A team approach is used to facilitate screening large populations. Each station requires approximately five working minutes for completion of seven tasks. The stations are referred to as Gross Motor, Fine Motor, Concepts, and Communications. Each operator (examiner) scores each child on a single sheet which moves from station to station with the child. In addition, social and affective behaviors are noted by every operator throughout the screening procedure.

The room arrangement includes a parent-observation area. Children can observe their parents and friends while participating in DIAL. Mothers

FIGURE 1

GROSS MOTOR

	INPUT			OUTPUT	
Items	Visual	Auditory	Haptic	Verbal	Motor
Throwing	X	X			X
Catching	X	X	X		X
Jumping	X	X			X
Hopping	X	X	X		X
Skipping	X	X			X
Standing Still	X	X			X
Balancing	X	X	X		X

FINE MOTOR

	INPUT			OUTPUT	
Items	Visual	Auditory	Haptic	Verbal	Motor
Matching	X	X	X		X
Building	X	X			X
Cutting	X	X			X
Copying Shapes	X	X			X
Copying Letters	X	X			X
Touching Fingers	X	X			X
Clapping Hands	X	X			X

CONCEPTS

	INPUT			OUTPUT	
Items	Visual	Auditory	Haptic	Verbal	Motor
Sorting Blocks	X	X			X
Naming Colors	X	X		X	X
Counting		X X		X	X X
Positioning		X			X
Following Directions		X			X
Identifying Concepts	X	X			X
Identifying Parts	X	X	X		X

COMMUNICATIONS

	INPUT			OUTPUT	
Items	Visual	Auditory	Haptic	Verbal	Motor
Articulating	X	X		X	
Remembering		X		X	
Naming Nouns & Verbs	X	X		X	
Coping		X		X	
Identifying Self Age Sex	X	X		X	
Classifying Foods		X		X	
Telling a Story	X	X		X	

FIGURE 2

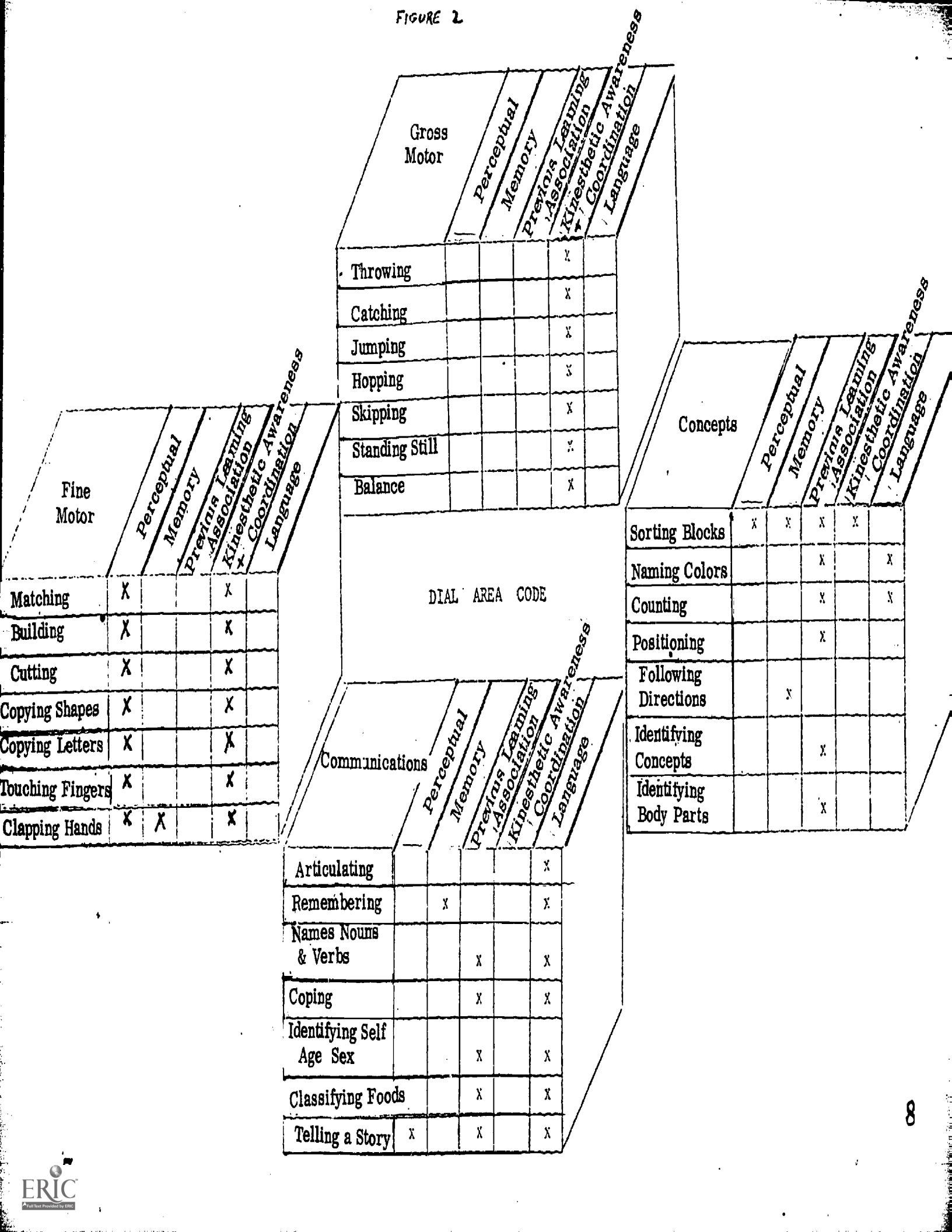
Gross Motor
Perceptual
Memory
Previous Learning
Association
Kinesthetic Awareness
Coordination
Language
Throwing
Catching
Jumping
Hopping
Skipping
Standing Still
Balance
Fine Motor
Matching
Building
Cutting
Copying Shapes
Copying Letters
Touching Fingers
Clapping Hands
Concepts
Sorting Blocks
Naming Colors
Counting
Positioning
Following Directions
Identifying Concepts
Identifying Body Parts
DIAL AREA CODE
Communications
Articulating
Remembering
Names Nouns & Verbs
Coping
Identifying Self Age Sex
Classifying Foods
Telling a Story

TABLE I

SUMMARY OF VARIOUS PLOTS SHOWING RELATIONSHIPS BETWEEN 1972 and 1973 SCORES

Figure	Item	Group	Correlation[3]	Slope	Intercept	Standard Error
1	Gross Motor	E	.436	.274	19.018	5.978
2	Gross Motor	C	.598	.350	18.367	3.642
3	Fine Motor	E	.494	.334	20.290	7.032
4	Fine Motor	C	.674	.391	19.200	3.752
5	Concepts	E	.428	.318	16.644	5.266
6	Concepts	C	.524	.161	20.109	2.214
7	Communication	E	.316	.159	26.374	6.971
8	Communication	C	.428	.187	26.918	2.406

[3]Correlation is between 1972 and 1973 scores.

can view their children's progress from a short distance and provide any moral or physical support needed.

1972

During 1972, a stratified sample design was constructed to balance selection of children on the basis of race, socio-economic status and demographic characteristics. Eight sites throughout the State of Illinois cooperated in conjunction with Title VI regional directors and coordinators. Specifications were established with each site requiring original score-sheets and data on a specific number of subjects. Approximately 4,400 children in the age range of 30 - 36 months were screened.

At no time was there any intention for DIAL to diagnose specific learning problems. Also, there was no plan to provide school districts or staff with specific suggestions, action, or DIAL scores for children screened. This action was taken to reduce contamination of the longitudinal validity study.

Both professional and paraprofessional staff served as DIAL operators. Careful monitoring and evaluation of DIAL teams consistently supported use of paraprofessionals as DIAL screeners.

Statistical analyses indicated that of the 124 items in the initial DIAL battery, 118 were statistically significant in regard to age. Children did better on the tasks as they grew older. The sex of a child was another important factor of performance. The mean score on females, on 15 selected items, exceeded the males. The trend suggested separate male and female norms.

A scoring cut off by sex and three month age segments reflected these developmental norms. In addition, a simple means of recording the child's response was developed whereby raw scores become scaled scores with relative ease. This scoring procedure results in a quick reporting system so the outcome of the screening can be determined in a matter of minutes.

Concurrent validity was established when a team of professionals spent three days to identify the same children (.92 agreement) identified by DIAL in three hours.

Items not statistically significant were removed from the DIAL battery. In addition, each DIAL item was analyzed in terms of input (how the task was presented to the child -- visually, auditorially, and/or haptically) and output (what was expected from the child -- verbally, and/or motorically) to aid in the identification of strong and weak modalities (Figure 1).

DIAL items were also analyzed in terms of the abilities they were intended to assess. The DIAL area code (Figure 2) gives this information. There are many DIAL items which are not unidimensional.

One of the incidental outcomes of the first year of Project DIAL research was a compilation of "Instruments for Screening of Prekindergarten Children" (Mardell & Goldenberg, 1972b) which provided educators with a description of the basic components of ninety evaluative instruments. The compilation indicates for each test the:

1. age range
2. depth (screening or diagnostic)

3. administration factors
4. modality of response
5. dimension(s) assessed
6. measurements requiring subjective
with/without child.

The intent of this document was to create an awareness of available instruments and their components rather than to assess their effectiveness. This document is available through ERIC.

1973

During 1973, a sample of 520 children, ages 2½ to 5½ years, was drawn from the 1972 pool of 4,400 DIAL screened children. DIAL items were reviewed to establish the conglomerate of skills necessary for acceptable classroom performance. Minor changes in the DIAL battery were made on the basis of statistical analyses and feedback from operators in the field. An auxillary weighted scoring system was designed by a statistician to separate the 85-90 percent of the population with no expected problem from the 10-15 percent who would be expected to have problems once in school. (Wick, 1973). The original and weighted scoring systems were applied to DIAL scoresheets and estimates of no risk-high risk were made.

The 520 children were retested with the revised DIAL. Half of the sample, randomly drawn, were part of the lowest 10-15 percent (experimental) and the other half were drawn from the top 90 percent (control). Correlations were based only on items used in both 1972 and 1973 versions of DIAL (Table 1).

Looking at Table 1, one will notice without exception correlations for the experimental group were less than that of the control group. This would be expected due to the constriction of range of scores within the groups. As the score range constricts, so does the correlation decrease. The communication area, being the most revised section, shows the lowest correlation.

All of the correlations are well into the range of significance, considering the severe restrictions of the scale. Possibly, if all the items had been included in the DIAL 1973 edition, there would, most likely, have been a test-retest correlation in the 0.65 range and higher.

Along with individual DIAL retesting, an external criterion of parent or teacher interview was applied to each subject's data profile. The questionnaire closely parallels the items in the DIAL battery.

A chi square statistic was used to test a hypothesis that the original categorization could not be predicted. Both null hypotheses were rejected beyond the .001 level of significance. Both parents and teachers were able to discriminate between the control and experimental groups in terms of development.

These results were limited, however, due to accuracy. Although the questionnaire results are statistically significant, they are not predictive. A prediction prescribes accuracy. The inaccuracy stems from a) possible errors in sample identification, b) error of measurement inadequacies of the interview design, parent or teacher information error, purposefully or inadvertently. In consideration of the above, it would be most inappropriate to use the interview as a replacement of the DIAL battery.

In addition to following 520 children, norms were reconstituted on a population of 3100 children, using one hundred randomly selected for each month of age from 36 to 66 months and controlled for sex so that male and female norms were established.

Conclusions of the 1973 study by Wick state that "The results are not unreliable. Reliable implies pre eliability implies that results occur in a random sort o Reason through the process involved in this project: Children (260 of them) were identified based on maturational lag from the 1972 testing. They, and another group (260) not identified as having a maturational lag, were retested in 1973 with the same battery twelve months later. If the reliability of the test were quite low, due to unreliable items or unreliability in the testers, the results would have blurred. That is, the groups identified as separate in 1972 ob y would not have appeared very distinct in 1973 -- if reliability was minimal. The 1973 groups were, actually, very distinct -- even after a full year since original identification. The evi in the technical section indicates that the battery is reliable."

1974

During 1974, 249 children from the 1972 sample were followed. The characteristics of these children are summarized in Table 2. As can be seen, 86 children were completing kindergarten and 163 were completing first grade. External criteria used were the Metropolitan Reading Readiness Test for kindergarteners and one of three achievement tests*

---

*Iowa Test of Basic Skills, Metropolitan Achievement Test, Stanford Achievement Test

TABLE 2.

SITE	POPULATION CHARACTERISTICS			# SOUGHT			# FOUND			% OF # SOUGHT	% OF TOTAL SAMPLE		
Joliet	Urban			144			135			94%	54%		
		Black	Nonblack		72	72		69	66			28%	26%
		Low	Nonlow		72	72		69	66			28%	26%
		Kgn	First		48	96		43	92			17%	37%
Lake County	Non-			72			71			99%	29%		
	Urban	Black	Nonblack		0	72		0	71			0%	29%
		Low	Nonlow		36	36		36	35			15%	14%
		Kgn	First		24	48		24	47			10%	19%
Elgin	Non-			72			43			60%	17%		
Freeport	Urban	Black	Nonblack		72	0		43	0			17%	0%
Posen-Robbins		Low	Nonlow		36	36		21	22			8%	9%
Quincy		Kgn	First		24	48		19	24			7%	10%
				288			249			86%	100%		

Table 3

Percent Agreement in

Scoring Among 16 DIAL Scorers

Area	Child Number	% Agreement
Gross Motor	1	.81
	2	.90
Fine Motor	1	.83
	2	.91
Concepts	1	.95
	2	.99
Communications	1	.91
	2	.97

It can be readily seen that the scoring of the first child in each of the four areas is not as reliable as the scoring of the second child in each area. This may be due to the fact that the first child was the younger child, still a year and a half away from the kindergarten, or due to the anxiety level of the operator and/or the scoring of the videotape itself.

In the Gross Motor area, the most unreliable items were scoring of "throwing","hopping", and "balance beam". After the manual directions are revised, it will be possible to establish if reliability can be significantly increased on these two items.

In the Fine Motor area, the scoring directions for drawing shapes and letters need revision. In the Concepts area, errors in scoring "front" and "back" may be avoided through a manual clarification. Otherwise, this area had particularly high reliability.

for first graders plus a teacher rating scale which covered social and affective behaviors. Multiple correlation with the test criterion measures using the DIAL scoresheet and weighted scale systems range from .47 to .60, all being significant. It is highly probable that these multiple correlations would be higher had the same test battery been used for all of the children. The correlations between DIAL and the teacher rating scale were not significant. Analysis of population dichotomies by race, SES, and demographic characteristics indicates no significant differences between these groups. Norms for boys and girls will, however, be maintained.

Experts in the fields of child development, special education, educational psychology, and related fields served as consultants to substantiate content validity.

Inter-rater reliability was established by means of viewing a videotape of children being screened. Table 3 shows the percent of agreement among sixteen scorers in each of the DIAL areas.

Conclusions

The development and refinement of testing procedures for identification of prekindergarten children with developmental delay for the past three years has substantiated the following:

1. It is possible to screen prekindergarten children and identify those children in need of further diagnostic evaluation.
2. There is a significant corelation between DIAL scores and readiness and achievement test scores.

3. There is a high percent of agreement among DIAL scorers.

4. Studies thus far only follow children through first grade. Longer studies are necessary to provide information about children who do not experience difficulty in school until third or fourth grade.

REFERENCES

Hall, J., Mardell, C., Wick, J. & Goldenberg, D. Final Report, Part I: Development and Refinement of DIAL. Resources in Education, May, 1974.

Mardell, C. & Goldenberg, D. Learning Disabilities/Early Childhood Research Project, Annual Report. Springfield, Illinois, Office of Superintendent of Public Instruction, 1972 (a).

Mardell, C. & Goldenberg, D. Instruments for Screening of Pre-kindergarten Children. Springfield, Illinois, Office of Superintendent of Public Instruction, 1972 (b).

Mardell, C. & Goldenberg, D. "For pre-kindergarten screening information: DIAL.", Journal or Learning Disabilities, March 1975.

Wick, J. et al. Validation and Normative Study of the DIAL Battery, Final Project Report to the Office of the Superintendent of Public Instruction, Evanston, Illinois, Northwestern University, 1973.