In studies conducted in the Baltimore City Public Schools by the Office of Pupil and Program Monitoring and Appraisal (OPPMA), pre-kindergarten pupils were found to be "growing scholastically". A program, Early School Admissions Program (ESAP), was based on nine objectives established to prepare pre-schoolers for "reading readiness" when they entered the kindergarten. To assess the success of the children and the program, a criterion-measure, based on Piagetian concepts, was constructed in cooperation with OPPMA and the Project Manager of ESAP. Reliability and validity of the Program Evaluation Measure (PEM) were established and found to be satisfactory, thereby, giving more credence to the growth indicators of the pupils. The profiles of the children (th.) were tested in October, November, December, January and May) demonstrated a significant pattern of achievement similarity as reflected by Dumas Model for measuring profile configurations. The Dumas Model for determining the similarity of profile configurations enables one to quantify the degree to which several profiles are shaped alike.

(Author)
INSTRUCTIONAL EFFECTIVENESS IN A PRE-KINDERGARTEN PROGRAM:
A PUPIL APPRAISAL AND PROGRAM EVALUATION MODEL

Edward N. Whitney, Ph.D.
Staff Director
Office of Pupil and Program Monitoring and Appraisal

Ernestine M. Reid, Ph.D.
Specialist
Office of Pupil and Program Monitoring and Appraisal

John L. Crew, Ph.D.
Deputy Superintendent
Center for Planning, Research and Evaluation
ABSTRACT

In studies conducted in the Baltimore City Public Schools by the Office of Pupil and Program Monitoring and Appraisal (OPPMA), pre-kindergarten pupils were found to be “growing scholastically”. A program, Early School Admissions Program (ESAP) was based on nine objectives established to prepare pre-schoolers for “reading readiness” when they entered the kindergarten. To assess the success of the children and the program, a criterion-measure, based on Piagetian concepts, was constructed in cooperation with OPPMA and the Project Manager of ESAP.

Reliability and validity of the Program Evaluation Measure (PEM) were established and found to be satisfactory, thereby, giving more credence to the growth indicators of the pupils. The profiles of the children (they were tested in October, November, December, January and May) demonstrated a significant pattern of achievement similarity as reflected by Dumas Model for measuring profile configurations. The Dumas Model for determining the similarity of profile configurations enables one to quantify the degree to which several profiles are shaped alike.
Instructional Effectiveness In A Pre-Kindergarten Program: A Pupil Appraisal and Program Evaluation Model

Edward N. Whitney, Ph.D.  Ernestine M. Reid, Ph.D.
John L. Crew, Sr., Ph.D.

INTRODUCTION

One of the prime functions for the reorganization of the Baltimore City Public Schools System last June was to improve the instructional program for all pupils. For that purpose, the Office of Pupil and Program Monitoring and Appraisal was established, with its primary goal being, to provide measurement and evaluation techniques which would be most appropriate for instructional programs. Over the years, the evaluation of the Early Admissions Program consisted primarily of pre and post intelligence and achievement tests, practically all commercially acquired. This procedure, though widely used, has left gaps in the kind of instructional feedback that could be given because of one or several of the following: a) this technique often omitted the characteristics of the pupils themselves; b) sociological characteristics which would have taken into account variables such as one parent, two parents, relative income or degree of educational deprivation of the parents when the child was first exposed to the program; c) the measurement of outcomes in terms of global standardized tests which may or may not have been related to instructional programs; d) strategies that apparently did not provide instructional personnel interim ideas related to individual or small group pupil progress, therefore, programs could not be adjusted or self-correcting as the school year progressed, thus enabling instructional personnel to effectively deal with each pupil.

STATEMENT AND PROBLEM

For all of the above reasons, Baltimore City attempted to develop a model procedure of pupil appraisal and program monitoring which would, in fact, correct the deficiencies of the evaluation strategies noted above. The problem was to develop an instruction-based measuring system which, in part, embraced learning constructs delineated by Jean Piaget. Integral to the development of such a system is its efficiency as an assessment device that is not dependent on traditional standardized tests. In addition, such studies as one by Yawkeye (1971) relating to the Early School Admissions Program (ESAP) Objectives found that pre-kindergarten children, when reinforced with certain numerical tasks, such as conservation of numbers, not only were successful in developing in-number concepts, but in cognitive development as well.
STATEMENT AND PROBLEM (continued)

The Baltimore City study found that Early Admissions pupils improved in their ability to verbalize, as reflected on a criterion measure, regardless of economic backgrounds and/or Binet test scores. This tends to discount findings of Audren Little in her study conducted to assess cognitive development in young children.

OPPMA aspires toward a model evaluation for pre-kindergarten education which could be applied to other programs in grades beyond pre-kindergarten and thus demonstrate methods to improve the effectiveness of school programs in general, but in particular, programs affecting inner-city pupils.

STUDY HYPOTHESES

Hypothesis One
An intervention program such as Early School Admissions Program (ESAP) will impact pupils, as reflected on specified tests, (i.e., the Pupil Evaluation Measure), to the degree that achievement characteristics will be statistically non-distinguishable as a function of low income pupils coming from homes having both parents, one parent or having a guardian.

Hypothesis Two
The periodic administering of the criterion measure (PEM), on a monthly basis, or longer time-interval, will demonstrate significantly at the .05 level, a continuity of profile similarity of pupil increments or learning, uniquely related to each of the nine program objectives.

PROCEDURES AND METHODOLOGY

During October 1973, the total population (1,260) of the Early School Admissions Program was administered the instrument referred to above, the Pupil Evaluation Measure (PEM). The PEM had been developed by OPPMA in close collaboration and cooperation with the Project Manager of the Early School Admissions Program.

The design for this study is characterized as a time-series testing of the total population, without a control group. The locally developed criterion device reflected the program objectives and required pupils to perform tasks which, behaviorally, conformed to aspects of Piagetian constructs of cognitive development.
PROCEDURE AND METHODOLOGY (continued)

The program objectives, (see Table 2), transcend nine different instructional disciplines, each showing a general degree for which cognition, as defined by Piaget, has developed in the individual pupil. A child’s performance on the PEM criterion measure has been structured so that his relative performance across each cognitive dimension (i.e., concrete, semi-concrete, abstract and semi-abstract) can be quantified. Each dimension has contributed evenly to the items of the test, and each have a maximum proportionate value of twenty-five percent (25%).

Assessment of the stability of perceptions among the monitor while administering the locally developed criterion measure, and some measure of the reliability and validity of the criterion measure (PEM) was needed.

RESULTS

The Early School Admissions Program (ESAP) was a pre-kindergarten program conducted for the purpose of developing skills that would prepare 4 year-olds for “reading readiness” upon their entrance into kindergarten. The pupils in the program have been identified as those 4 year-olds living within an area defined as inner-city and disadvantaged. Family income, as established by the Office of Economic Opportunity, was one determinant of pupil eligibility, as were the limited educational opportunities of adult members of the family, dependency on public assistance and varied problems plaguing the family.

Using two tests, the Stanford Binet (an intelligence quotient measure) and a locally constructed criterion (objective) reference measure, Pupil Evaluation Measure (PEM), OPPMA set out to observe each child’s achievement over the school year.

The instrument has been checked to determine its reliability through a procedure involving relative agreement among monitors on the PEM instrument. The inter-monitor reliability coefficient of the instrument was in excess of .85, thus knowledge regarding the instrument’s administering stability was known.

The consistency in which the pupils have performed on this test has been determined through a reliability computational procedure known as the Kuder 21, Equation.
RESULTS (continued)

Interest in reliability center on how the boys and girls consistently performed for each month the test was administered. As such, five measures of reliability were determined, and the following Table reflects corresponding computational reliabilities for each of these months.

<table>
<thead>
<tr>
<th>MONTHS</th>
<th>RELIABILITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>October</td>
<td>0.6748215</td>
</tr>
<tr>
<td>November</td>
<td>0.5944445</td>
</tr>
<tr>
<td>December</td>
<td>0.302100</td>
</tr>
<tr>
<td>January</td>
<td>0.8117085</td>
</tr>
<tr>
<td>May</td>
<td>0.8576863</td>
</tr>
<tr>
<td>Average</td>
<td>.74</td>
</tr>
</tbody>
</table>

The overall reliability average computed to .74 which reflects a moderately high reliability coefficient.

Validity has been determined by comparing sub-sections of the Stanford Binet Intelligence Test (Short Form) with corresponding sections of the Pupil Evaluation Measure (PEM). Validity is designed as the degree to which the scale measures the concept on which the study purports. As such, validity is the crucial test of a scale. Logical validity had been established primarily as a result of the manner in which the test has evolved. Previously, it has been stated that the Project Manager of the Early School Admissions Program was intrically involved in the development of this instrument. The items of the Pupil Evaluation Measure (PEM) reflect the nine program objectives of the ESAP.

Construct validity has been assumed through a process most commonly identified as the Spearman Rank Order Correlation Procedure. In a publication by Haven, Rogers, and Lipmann, it is suggested that validity can be estimated by observing the comparability of performance of one instrument to performances by the Subjects on an existing instrument.
RESULTS (continued)

“Sometimes, several independent criteria are available
to measure a single dimension, and the researcher wishes
to devise a simpler measure to assess the dimension, so,
he may relate the newly constructed measures to these
criteria.”

In assessing the Pupil Evaluation Measure scores ranked by school, and comparing them with the Stanford Binet rank order scores, there was found to be high validity between the two instruments. Data from both instruments were collected in October when the Stanford Binet was administered as a pretest and the PEM was administered as a measure of the degree to which the program objectives were met. PEM was administered each month thereafter, other than February, March and April, due to system-wide reasons.

The items in PEM and the sub-tests of the Stanford Binet were observed carefully and found to be related. Chein defines construct validity as the degree to which phenomena can be observed on separate criterion designed to assess similar traits, conditions or phenomena. The construct validity coefficient, so designed, using these instruments during the period between October and May, was moderately high at the .84 level.

Based on previous reports in the literature and general positions espoused by many educators, the data was examined for six months from the first administration of PEM in October 1973 to determine if there were relative similarities of pupil performance on the instrument in terms of specified sociological and demographic variables. The variables examined were one parent in the home, two parents in the home, guardian, aid to families with dependent children and non-aid to families with dependent children.

The criticalness of these variables are accentuated by the fact that insight can be more realistically determined as to the relative effectiveness of intervention programs such as ESAP are having on a more comprehensive basis.
RESULTS (continued)

Using information gathered from Pupil Personal History Forms, each child was categorized according to his socio-economic background. The categories and test results, according to each of the nine objectives, follow:

A. Pupils having in the home:

1. Both Parents (B)
2. One Parent (O)
3. Guardian Parents (G)

B. Receiving Aid To Dependent Children:

1. No (N)
2. Yes (Y)

C. Schools with:

1. All Black Students
2. All White Students
3. Mixed Students

After many observations, evaluations and assessments, the analysis proved to be rewarding, "our children were growing". Some were growing faster than others and some children achieved higher levels than others, but they were achieving, each at his own rate. The results indicated each child to have weaknesses in some skill areas and strengths in others. This indicated the possibility of building on strengths to improve weaknesses.

The evaluation of each objective of the program was done so as to follow the child's development in his ability to transfer from concrete reasoning to the level of abstract reasoning. Data were included to demonstrate achievement by child, by class, by school, by program and by pupil background.
The purpose of the evaluation by objective was to be able to determine the levels of development demonstrated by pupils. A determination could be made as to areas where a pupil may need heavy concentration in order to help him develop in the acquisition of skills. These skills included classification, verbal, visual and auditory perceptual, motor perception, recognition and use of art media and the ability to reason and think through problems. Another “must area” was self-concept, children need to “feel good” about themselves to be successful. The scoring was as follows:

- 0 - 24 - demonstrated the pupil working at concrete level
- 25 - 49 - demonstrated the pupil at the semi-concrete level
- 50 - 74 - demonstrated the pupil at the semi-abstract level
- 75 - 100 - demonstrated the pupil achieving at the abstract level

The following data indicate the number of children in the Early School Admissions Program who were below or above the semi-concrete/semi-abstract stage of cognitive development in October 1973 and how they appeared in June of 1974. The summaries indicate, by objective, those above the 50% levels.
ILLUSTRATION OF ACHIEVEMENT PATTERNS
FOR FIVE SEPARATE MONTHS FOR PUPILS
COMING FROM HOMES OF 2P, 1P, G, ADC, NON-ADC

TABLE 2

Each pupil will demonstrate the ability to identify, name and describe a variety of concrete
objects including people, animals and toys.

<table>
<thead>
<tr>
<th>Time Tested</th>
<th>October</th>
<th>June</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 50%</td>
<td>755</td>
<td>230</td>
</tr>
<tr>
<td>Above 50%</td>
<td>262</td>
<td>845</td>
</tr>
</tbody>
</table>

Each child will demonstrate individual development of furthering his language ability in
conversation, self-expression and relating experiences.

<table>
<thead>
<tr>
<th>Time Tested</th>
<th>October</th>
<th>June</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 50%</td>
<td>632</td>
<td>185</td>
</tr>
<tr>
<td>Above 50%</td>
<td>294</td>
<td>823</td>
</tr>
</tbody>
</table>

KEY

- **ONE PARENT**
- **BOTH PARENTS**
- **GUARDIAN**
- **NON-ADC**
- **ADC**
ILLUSTRATION OF ACHIEVEMENT PATTERNS
FOR FIVE SEPARATE MONTHS FOR PUPILS
COMING FROM HOMES OF 2P, 1P, G, ADC, NON-ADC

TABLE 2

<table>
<thead>
<tr>
<th>OBJECTIVE 3</th>
<th>OCT</th>
<th>NOV</th>
<th>DEC</th>
<th>JAN</th>
<th>FEB</th>
<th>MAR</th>
<th>APR</th>
<th>MAY</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIME TESTED</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>October</td>
<td>452</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>June</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Each child will use problem situations to develop thinking and reasoning skills.

<table>
<thead>
<tr>
<th>OBJECTIVE 4</th>
<th>OCT</th>
<th>NOV</th>
<th>DEC</th>
<th>JAN</th>
<th>FEB</th>
<th>MAR</th>
<th>APR</th>
<th>MAY</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIME TESTED</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>October</td>
<td>933</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>June</td>
<td>126</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Each pupil will develop classification skills and will be able to group concrete objects by form, size, color and function.
ILLUSTRATION OF ACHIEVEMENT PATTERNS
FOR FIVE SEPARATE MONTHS FOR PUPILS
COMING FROM HOMES OF 2P, 1P, G, ADC, NON-ADC

TABLE 2

<table>
<thead>
<tr>
<th>Objective 5</th>
<th>October</th>
<th>June</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 50%</td>
<td>927</td>
<td>141</td>
</tr>
<tr>
<td>Above 50%</td>
<td>297</td>
<td>687</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Objective 6</th>
<th>October</th>
<th>June</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 50%</td>
<td>897</td>
<td>218</td>
</tr>
<tr>
<td>Above 50%</td>
<td>524</td>
<td>537</td>
</tr>
</tbody>
</table>
ILLUSTRATION OF ACHIEVEMENT PATTERNS
FOR FIVE SEPARATE MONTHS FOR PUPILS
COMING FROM HOMES OF 2P, 1P, G, ADC, NON-ADC

TABLE 2

Each pupil will develop quantitative skills by being able to identify, match and compare sets of objects, count, and develop basic understanding in addition and subtraction, as reflected.

<table>
<thead>
<tr>
<th></th>
<th>Below 50%</th>
<th>Above 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>October</td>
<td>798</td>
<td>224</td>
</tr>
<tr>
<td>June</td>
<td>301</td>
<td>772</td>
</tr>
</tbody>
</table>

Each pupil will develop wholesome feelings of self as reflected by his ability.

<table>
<thead>
<tr>
<th></th>
<th>Below 50%</th>
<th>Above 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>October</td>
<td>934</td>
<td>341</td>
</tr>
<tr>
<td>June</td>
<td>174</td>
<td>776</td>
</tr>
</tbody>
</table>

KEY
- ONE PARENT
- BOTH PARENTS
- GUARDIAN
- NON-ADC
- ADC
Each pupil will have opportunities to increase those abilities that allow him to use a variety of art media and many art forms.

<table>
<thead>
<tr>
<th>Time Tested</th>
<th>Below 50%</th>
<th>Above 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>October</td>
<td>1072</td>
<td>661</td>
</tr>
<tr>
<td>June</td>
<td>2</td>
<td>345</td>
</tr>
</tbody>
</table>

The data demonstrated the reading readiness of boys and girls in the Early School Admissions Program as they entered kindergarten in September of 1974. More than 77% of the pupils appeared ready for the kindergarten program and beyond.

Included within this report are additional charts explaining other aspects of the studies (racial, social, economic). The in-depth analysis included herein proves the validity and reliability, as well as the value of this report.
MEASURES OF PROFILE SIMILARITIES OF PUPIL ACHIEVEMENT
AS OBSERVED OVER A FIVE-MONTH-TIME SPAN

Profile analysis, such as defined by Dumas, is the study of profile similarity between test scores, and it has a position of real value in the research field. Probably no phase of educational administration has received more criticism than the evaluation of performance. Profile analysis and the application of one of the techniques currently available will permit far better evaluation of efficiency than any of the old rating scale techniques.

The similarity of continuity of growth, as reflected on the criterion measure, indicated a pattern of similarity that was significant beyond the .05 level. The Dumas Model of Profile Similarity computed to a value of .745 overall for the school year ending June 1974. Such analyses would tend to reflect program impact that has generally resulted in pupils, affected by this program, demonstrating differential achievement trends across the nine program objectives, but more important, demonstrating continuity of achievement during the school year. Table 3 pictures even better the profile similarities of achievement across the five months of testing.
Table 3

MEASURES OF PROFILE SIMILARITIES OF PUL.'L ACHIEVEMENT
AS OBSERVED OVER A FIVE-MONTH-TIME SPAN

According to the data accumulated, we have a moderate to high level of profile similarity.

\[ r_{ps} = .745; p < .05 \]
DISCUSSION

The importance of the problem is accentuated by the fact that of the numerous in-house and outside evaluators variously involved in former evaluations, the product emphasis apparently has been more in terms of program implementation, the operational aspects, and subjective reactions on the part of evaluators basing and concluding programmatic impact as a function of anecdotal references. Typical of the manner in which evaluations of the Early School Admissions Program has been affected or implemented, one finds in the official report of the 1966-67 ESEA Early School Admissions Program, on page 4, one notes:

"It is difficult, on the basis of a short observation, to note tangible evidence of possible effect of a program on the children involved. There were, however, some differences in varying degrees, noted by the three evaluators. During each brief school visit, the team members divided their efforts as follows: One member of the team interviewed parents in a group situation; while another, when possible, talked with children now in kindergarten and with Non-Early Admission children who were in the program last year; and the third member observed and talked with children and interviewed parents at each center visit.

Findings were then later compared school-by-school with the evaluation compared by each center at the close of the school year. However, the reported findings are not related to the individual schools. The material was written after discussing procedures. They were then collated, typed, and reacted to by each individual team member, prior to preparing the final report. The evaluators feel that the final report is a team effort and represents their best combined judgment within the limitation imposed.

The validity of the report thus rests on the impressions received and reacted to on the basis of the experience and philosophy of the evaluators, as well as on the accuracy of the information provided by the school personnel."
DISCUSSION (continued)

Serious concerns are in order if the preceding quote typifies evaluational procedures previously used on the ESAP. Another rather extensive evaluation submitted by a former Research Director, one finds extensive information on demographic, social, economic and medical type information being presented, but programmatic impact on the part of the pupils is never specifically delineated or attempted to be delineated other than generalized intuitive reactions on the part of the evaluator.

Obviously, insight is needed relative to whether or not the program itself is affecting the boys and girls and this should be, as thoroughly as possible, emperically substantiated. This has been the underlying basis of the current report which essentially attempts to delineate a multi-variant approach in terms of how girls and boys are being affected and impacted by the Early School Admissions Program operating during the 1973-74 school year in Baltimore City. Nine specific objectives have been delineated, and the relative impact and inter-relationship of these different objectives has also been observed and documented.

In addition, documentation of the program impact on pupils has been coupled with a construct which relates to a Piagetian definition of cognitive development, on several dimensions of cognitive growth on the part of pupils. Efforts have been made to make associations as far as these nine objectives are concerned to specified aspects of the Piagetian concept of cognitive development. Some intriguing issues, as addressed by this study, were determining 1) whether pupils are being affected by an intervention program such as the Early School Admissions Program; 2) to determine the various ways or dimensions in which such effect is actually taking place with the children; 3) to determine whether or not specified sociological differences can be determined as a function of boys and girls being exposed to this program; 4) to determine the longitudinal implications as a function of observing boys and girls, who were former Early School Admissions pupils, who may now be in the kindergarten or other grades, and determine if their subsequent achievement pattern indicates advantage over boys and girls who may not have had either an Early School Admissions Program experience or any pre-kindergarten experience; 5) the manner in which mechanisms can be set up sensitive enough to determine to what extent validation procedures can be used for varifying the fact of observed discrepancies of different observations.
DISCUSSION (continued)

It is extremely important that educators acquire the insight and sensitivity as public school systems increasingly progress toward individualization, and maximally adapt school systems to accommodate the unique needs of individual pupils. Public school officials should be cognizant of the extent in which instructional experiences are broad and comprehensive and can be delineated according to some pedagogic philosophy or educational learning construct. Program impact and school experiences can be observed more practically in a manner which is most relevant to the individual child. The approach for assessing the effects of this program on the various boys and girls in the project has been established to the degree that we now see or have seen how boys and girls achieve, and while their relative growth rates may not be the exact level across all objectives, achievement has been observed across all objectives.

One of the major ramifications of this study relates to the manner in which the individual pupil may be observed in a comprehensive manner which allows continuous observation on that individual over several variables, factors, or objectives. A possible use on other grade levels of the school system are rather apparent as far as a similar model can be devised that allows the individual pupil to be observed or assessed across several related factors in one particular skill area.

The approach for assessing the effects of this program on the various boys and girls in the project has been established to the degree that we now see or have seen how boys and girls achieve, and while their relative growth rates may not be the exact levels across all objectives, achievement has been observed across all objectives.

SUMMARY

The Office of Pupil and Program Monitoring and Appraisal (OPPMA) has found a means of evaluation that may have far-reaching effects on the educational system of the Baltimore City Public Schools. Discovering learning patterns at the pre-kindergarten level can enable educators to prepare programs that may build a more effective educational process within the system. More children can be successful when the rate of learning is recognized early.
BIBLIOGRAPHY


Little, Audren. "A Longitudinal Study of Cognitive Development In Young Children." Child Development. 43:3 September 1972. 1024-34.


BIBLIOGRAPHY (continued)

