DOCUMENT RESUME

ED 054 511 24 CG 006 652
AUTHOR Thomas, Alexander
TITLE Preschool Behavioral Style and Later Academic Achievement. Final Report.
INSTITUTION New York Univ., N.Y. Medical Center.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
BUREAU NO BR-8-0663
PUB DATE 15 Jul 71
GRANT OEG-0-9-420663-3705(010)
NOTE 78p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Academic Achievement; Affective Behavior; *Behavior Patterns; *Behavior Problems; *Elementary School Students; Individual Characteristics; *Preschool Children; Student Behavior; Task Performance

ABSTRACT

The major aim of this study was to investigate the relationship between aspects of a child's behavioral style, specifically his temperament and task orientation, in the preschool age period and his level of academic achievement in elementary school. The relationship between behavior problems in the preschool years and later academic achievement was also studied. Results indicated that (1) the temperamental characteristics on non-adaptability and withdrawal showed the most significant correlations (negative) with academic achievement; (2) children with signs of the difficult child did not appear at risk for academic underachievement; (3) similarly, the characteristics of high activity and distractibility did not show any striking correlations with academic achievement; and (4) qualitative analysis of the behavior problem cases revealed a tendency for the low academic achievers to come from homes in which there was severe intrafamilial stress. Recommendations include that teachers be familiar with the characteristics of the slowly adaptable and withdrawn child and with the approaches desirable to maxamize his learning and school achievement. In addition, teachers should be aware that characteristics which are obviously troublesome in the school setting do not necessarily predict unfavorable academic achievement. (Author/TA)

PA 24

FINAL REPORT

Project No. 8-0663

Grant No. OEG-0-9-420663-3705(010)

PRESCHOOL BEHAVIORAL STYLE AND LATER ACADEMIC ACHIEVEMENT

Alexander Thomas, M. D.
New York University Medical Center
550 First Avenue
New York, New York 10016

July 15, 1971

FINAL REPORT

Project No. 8-0663

Grant No. OEG-0-9-420663-3705(010)

PRESCHOOL BEHAVIORAL STYLE AND LATER ACADEMIC ACHIEVEMENT

Alexander Thomas, M. D.
New York University Medical Center
550 First Avenue
New York, New York 10016

July 30, 1971

The research reported herein was performed pursuant to a grant with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

Table of Contents

Page

TABLES continued:

INTRODUCTION

The major aim of this study was to investigate the relationship between aspects of a child's behavioral style, specifically his temperament and task orientation, in the preschool age period and his level of academic achievement in elementary school. In addition, we also studied the relationship between behavior problems in the preschool or early school years and later academic achievement. The subjects in the study were 136 children of middle- and upper-middle-class background whom we have followed longitudinally from their first few months of life.

Behavioral Style

Behavioral style is a general term referring to how a child behaves. It differs from ability, which considers the what and how well of behavior, and from motivation, which seeks to account for why a person does what he is doing. Guilford (1959) and Cattell (1950) are other workers who have analysed behavior in terms of the how, the what and the why.

In examining behavioral style, our primary concern is with the way in which an individual behaves. Two children may each eat skillfully or throw a ball with accuracy and have the same motives for so doing. Yet, they may differ with respect to the intensity with which they act, the rate at which they move, the mood they express, the readiness with which they shift to a new activity and the ease with which they approach new toys, situations or playmates.

An individual's behavioral style is not immutable. As is the case with any other characteristic of the organism, its features can undergo a developmental course that will be significantly affected by environmental circumstances. In this respect it is not different from height, weight or intellectual competence and, as is the case for these and other characteristics of the individual, the pattern initially identified in the young child may remain relatively unchanged by environmental influences or it may be reinforced,

heightened, diminished or otherwise modified during the developmental course.

In recent decades, ability and motivation have been the main focus of behavioral studies by investigators in the field of child development and child psychiatry. Educators, on the whole, have concentrated on studying the relationship of specific features of ability (intellectual level, cognition, perception, etc.) to learning. While these studies have been extremely fruitful, examinations of the effects of motivation and psychodynamic factors on learning have been largely inconclusive (Raph, et al., 1966).

Studies of behavioral style, by contrast, have been until recently much fewer in number. Those done by workers in the fields of child development and psychiatry were, in general, limited, fragmentary and without long-term follow-up, and there was no attempt to correlate behavioral style with educational achievement. Thus, they have not provided a basis for systematic and comprehensive understanding of this feature of performance. Furthermore, in the field of educational research, studies of the influence of behavioral style on learning have been characterized by their absence.

Temperament as an Aspect of Behavioral Style

Since 1956, the principal investigator in this study and his colleagues have been engaged in a number of studies concerned with the systematic delineation and categorization of specific features of behavioral style in early childhood and the influence of these characteristics on normal and deviant psychological development. The children studied have included the 136 youngsters used in this project; children from Puerto Rican working-class families living in New York City; children born prematurely; mentally retarded children; and children with congenital rubella. In all these samples, it has been possible to identify and categorize quantitatively similar features of behavioral style. This work has progressed against a background of increasing attention to the study of various features of behavioral style by

a number of other investigators in the past 15 years (Kagan and Moss, 1962; Murphy, et al., 1962; Escalona and Heider, 1959; Bridger and Birns, 1963).

We have concentrated, in particular, on certain aspects of behavioral style that could be identified in the first few months of life and also at subsequent age periods and which we have called temperament.

Temperament is a phenomenologic term, with no inferences as to genetic, somatologic, endocrine or environmental etiologies. It describes the characteristic tempo, rhythmicity, adaptability, energy expenditure, mood and focus of attention of a child, independently of the content of any specific behavior. We have used nine categories of reactivity within which to subsume temperamental attributes. A child's temperamental organization thus represents his characteristic mode of functioning with respect to these aspects of his behavior.

Definitions of the temperamental characteristics and their ratings used in this and our other studies are as follows:

1. Activity level describes the motor component in a child's functioning and the diurnal proportion of active and inactive periods. Protocol data on motility during eating, playing, dressing, bathing and handling, as well as information concerning the sleep-wake cycle, reaching, crawling and walking are used in scoring this category. A child may be rated as having a high, moderate or low activity level.

2. Rhythmicity (regularity) describes the predictability or unpredictability in the timing of such functions as the sleep-wake cycle, hunger, feeding pattern and elimination schedule. A child may be rated as being regular, variable or irregular.

3. Approach-Withdrawal is defined as the nature of a child's response to a new stimulus, such as an unfamiliar person, food or toy. The scores for this characteristic are approaches, variable and withdraws.

4. Adaptability describes the child's responses to new or altered situations over a period of time. The concern here is not with the nature of the initial responses, but with the speed and ease with which they may be modified in desired directions. A child may be either adaptive, variable or nonadaptive.

5. Quality of mood refers to the amount of pleasant, joyful and friendly behavior a child displays as contrasted with unpleasant, unfriendly behavior and crying. A child's mood may be positive, variable or negative.

6. Intensity of reaction is defined as the energy level of a child's response, irrespective of its quality or direction. The ratings for this characteristic are intense, variable or mild.

7. Threshold of responsiveness refers to the intensity level of stimulation required to evoke a discernible response by the child to sensory stimuli, environmental objects and social contacts. A youngster may have a high, moderate or low threshold of responsiveness.

8. Distractibility is based on the effectiveness of extraneous environmental stimuli in interfering with or altering the direction of the child's ongoing behavior. A child may be rated as distractible, variable or nondistractible.

9. Attention span and persistence describe the length of time a particular activity is pursued and the continuation of an activity in the face of obstacles to maintaining the activity direction. The ratings for this category are persistent, variable and nonpersistent.

We have found that the most significant implications of the temperamental characteristics derive not from an examination of the separate categories but from a consideration of clusters of traits. Such clusters, or sets of attributes, have been found to be related to the development of behavior disorders and, in school-age youngsters, to academic performance.

One commonly found temperamental constellation is comprised of regularity, positive approach responses to new stimuli, easy adaptability to changes and preponderance of positive mood of mild to moderate intensity. A child with these temperamental characteristics develops regular sleep and feeding schedules easily, takes to most new foods at once, smiles at strangers, adapts quickly to a new school, accepts most frustrations with a minimum of fuss and learns the rules of new games rapidly. Such a youngster is aptly called the "easy child" and is usually a joy to his parents, pediatrician and teachers. The easy child generally adapts to the demands for socialization with little or no stress and confronts his parents with few, if any, problems in routine handling.

At the opposite end of the temperamental spectrum is the child with irregularity in biological functions, predominantly withdrawal responses to new stimuli, nonadaptability or slow adaptability to change, negative mood and preponderantly intense reactions. These five temperamental characteristics comprise what we have called the "difficult child" syndrome. The difficult child manifests irregular sleep and feeding patterns, slow acceptance of new foods, prolonged adjustment periods to new routines and frequent and loud periods of crying. His laughter, too, is characteristically loud. Mothers find such children hard to care for. They are not easy to feed, put to sleep, bathe or dress. New places, unaccustomed activities and strange faces all may produce initial responses of loud protest or crying. Frustration generally produces a violent tantrum.

We found in the New York Longitudinal Study and in our behavioral study of mentally retarded children

that youngsters with the cluster of temperamental traits characteristic of the difficult child accounted for a significantly high proportion of the behavior problem groups, well beyond their representation in each study sample. The specifically stressful demands for these difficult children were shown to be those of socialization, the demands for altering spontaneous responses and patterns to conform to the rules of living of the family, school or peer group. (Once these children do learn the rules, however, they usually function easily, consistently and energetically.) As a result of their temperamental characteristics, these children make demands on their parents for particularly consistent and objective handling if maladaptive temperament-environment interactions are to be avoided. If new demands are presented inconsistently, impatiently or punitively, effective changes in their behavior become stressful and even impossible. Negativism is a not infrequent outcome of such suboptimal parental functioning.

We have found no evidence that the parents of difficult children are essentially different from other parents, nor do our studies suggest that the temperamental characteristics of the youngsters are caused by their parents. Yet, the problems of managing a difficult child often highlight a parent's individual reaction to stress. The same parents who are relaxed and consistent with an easy child may become resentful, guilty or helpless with a difficult child, depending on their own personality structures. Other parents, by contrast, who do not feel guilty or put-down by the child's behavior may learn to enjoy the vigor, lustiness and "stubbornness" of a difficult youngster.

Another important temperamental constellation combines negative responses of mild to moderate intensity to new stimuli with slow adaptability after repeated contact. Children with this pattern differ from the difficult youngsters in that their withdrawal from the new is quiet rather than loud. They also usually do not have the irregularity of function, frequent negative mood expression and intense reactions of the difficult infant. The mildly expressed withdrawal from the new is typically

seen with the first encounter with a new food, a new person, a new place or a new task. If given the opportunity to re-experience these new situations without pressure, such a child gradually comes to show quiet and positive interest and involvement. This characteristic sequence of responses has suggested "slow to warm up" as an apt if inelegant appellation for these children. A key issue in their development is whether parents and teachers allow them to make an adaptation to the new at their own tempo or insist on immediate positive involvement, something which is difficult or impossible for them.

Of special interest, too, has been the qualitative study of the characteristic distractibility. Though the protocols usually show only a small number of scorable items in this category, it has been possible to trace the continuity of this trait in several children in different situations and at different ages. In these cases, it has become evident that the characteristic was influential in the pattern of the child's responses to feeding practices, toilet-training procedures, play situations and demands for various types of learning.

Task Orientation as an Aspect of Behavioral Style

An additional aspect of behavioral style has been defined and rated on the basis of a child's responses to the demands embodied in the test items of a standard psychometric test. This attribute, called task orientation, defines the way in which a child approaches a task when other choices for functioning are present. Within this definition, task orientation is rated independently of the child's actual level of performance of the task. In other words, the level of success or failure on the task does not influence the rating of task orientation.

A number of other investigators have called attention to the possible significance for a child's school functioning of behavioral style characteristics which may be similar to our category of task orientation. Kagan and Moss (1962), utilizing the data of the Fels Institute longitudinal study, have developed a method for rating behaviors considered to reflect achievement striving. The measures employed included estimates of degree of

independence and efforts at task mastery. The ratings at various age periods were studied, and they concluded that although achievement behavior during the first three years of life was essentially unrelated to later performance, achievement efforts in elementary school were predictive of adult achievement performances.

Crandall and her coworkers (Crandall, et al., 1962; 1965) found that measures of a sense of environmental control, a variable which appears significantly related to school achievement (Coleman, et al., 1966), were highly associated with the amount of time spent in intellectual activities during free play.

As with temperament, we have been able to categorize the various children we have studied on the aspect of their behavioral style which we have called task orientation. This has been done by scoring written descriptive accounts of a child's motoric behavior and verbatim reports of his verbalizations during the administration of the Stanford-Binet test. Each test item can be considered to represent a concrete demand on the child for specific task performance, and the total test to represent a standard series of such demands presented in sequence. A child's responses to these demands give a characteristic picture of his task orientation.

In describing a child's behavioral style in this area, the following categories were considered.

1. Work. The child actively engages in the task demanded of him, irrespective of the correctness of his responses.

2. Not Work. The child does not do what is demanded of him.

3. Spontaneous Extension. The child deals with the materials and elaborates on them.

4. Emotional Responsiveness. The child's response is accompanied by some overt display of emotions (laughing, smiling, crying, frowning).

5. Spontaneous Withdrawal. In between the presentation of test items, the child gets up and leaves the testing situation.

The "Not-work" responses can be further subdivided as follows:

a. Motor Not Work. The child removes himself physically from the task or pushes the material away.

b. Passive Not Work. The child makes no detectable response to the demands made of him.

c. Verbal Not Work. The child states that he will not do what is requested of him.

d. Verbal Not Work-ability. The child says that he does not know how to do a task.

e. Motor Substitution. The child performs a motor task other than the one demanded of him (for example, he draws a square when asked to draw a circle).

f. Verbal Substitution. The child verbally substitutes an alternative task (for example, he says, "I don't want to make a bridge -- I'll make a tower.").

Temperament, Task Orientation and Academic Achievement

In previous projects of the New York Longitudinal Study we have demonstrated clear relationships between aspects of a child's temperamental organization in the preschool period and the development of behavior problems. We have also suggested that a youngster's task orientation may be implicated in the elaboration of psychiatric disorders (Thomas, et al., 1968; Hertzig, et al., 1968). These findings have enabled us to develop individualized programs of intervention for the amelioration of disturbed behavior in many children.

We believed that if an investigation of the relationship between temperament and task orientation to academic achievement resulted in a similarly clear delineation of specific correlations, it would become possible to enhance programs of individualized instruction aimed at maximizing a child's learning and academic achievement, by furthering our understanding of his characteristics and potentialities. Furthermore, clarifying the relationship between behavior problems in the preschool or early school years and subsequent academic achievement should increase our understanding of which types of symptoms and problems are most likely to interfere with a youngster's progress in school.

METHODS

Sample

The children in this study were from families who had participated in the original New York Longitudinal Study of behavioral development from its inception. They were of middle- or upper-middle-class background and when we had first contacted them were resident in New York City or one of its surrounding suburbs. Families were enrolled in the study either during the mother's pregnancy or immediately following the birth of the child. The cumulative collection of families was completed over a six-year period during which 85 families with 141 children were enrolled.

Of the study families followed until 1967, the majority (78 per cent) were Jewish, with some Catholic and Protestant families. There were one Negro and one Chinese family. Almost all parents were born in the United States. Forty per cent of the mothers and 60 per cent of the fathers had both college educations and post-graduate degrees, and only 9 per cent of the mothers and 8 per cent of the fathers had no college at all. With only 3 exceptions, all of the fathers worked either in one of the professions, or in business at a management or executive level. Eighty per cent of the mothers had occupations similar to the fathers, and the remaining 20 per cent had been employed as office workers or secretaries.

Of the original 141 children enrolled in the NYLS, 136 were followed until formal study was ended in 1967. In August, 1968, the families were contacted again in order to investigate the question of academic achievement. The difficulties encountered in obtaining data in specific areas are discussed below. Nevertheless, it is worth noting here that new information was obtained on 116 of the youngsters, 57 boys and 59 girls. Discrepancies in the number of children used for various statistical analyses result from the fact that not all data could be gotten for all the children due to the range of their ages, lack of compliance with our requests by some schools

and the impossibility of interviewing all the children because of their geographical distance from New York.

Since the sample in the original study (Thomas, et al., 1968) was gradually accumulated, the children's birth years vary. Table 1 summarizes the ages of the children as of December, 1969, and indicates the normal grade placement for youngsters that age in the New York City school system. It should be noted that if a child had been accelerated or had attended private nursery-kindergarten school, he may actually be one grade higher than normal placement would indicate. This table includes all the children for whom substantial school achievement data were available. It can be seen that the vast majority of them (74 per cent) were in or had completed sixth grade.

Data Available at Start of Study

Temperament. Data necessary to define the child's temperamental organization and to delineate his developmental course were obtained from several sources. At regular intervals, interviews were held with the parents during which they were queried about characteristics of the child's behavior in the routines of daily living, his responses to any changes in these routines or in his environment and his reactions to any special events or life situations. Each year, in addition, information was obtained from his teacher regarding the child's initial adaptation to the class and his overall functioning throughout the term.

Supplementary information was also obtained through direct observations of each child by staff members, at least once a year in school and during the administration of standard psychological tests when each youngster was 3 and 6 years of age.

While we have used all these sources of data in developing a qualitative definition of a child's temperamental attributes, we relied on only the information obtained from the parents in making our quantitative categorization of temperament in the preschool years. Only the parent interviews represented an economical

Table 1

Age Distribution of Study Children

Age as of Dec. 1969	Boys	Girls	Total	School Grade*
8:6 - 8:11	0	1	1	3rd
9:0 - 9:5	1	2	3	4th
9:6 - 9:11	1	2	3	
10:0 - 10:5	2	9	11	5th
10:6 - 10:11	7	5	12	
11:0 - 11:5	2	2	4	6 h
11:6 - 11:11	4	4	8	
12:0 - 12:5	9	2	11	7th
12:6 - 12:11	4	6	10	
13:0 - 13:5	5	10	15	8th
13:6 - 13:11	15	10	25	
14:0 - 14:5	7	6	13	9th
Total	57	59	116	

*Normal grade placement for age in N.Y.C. Public Schools -- if accelerated or attended private nursery-kindergarten, child could be one grade higher.

source of detailed longitudinal behavioral information on a sufficiently large sample of children because of their continuous direct observation of the child. This could otherwise have been duplicated only by an investigator living in the home. The crucial question, whether the parental reports represented valid reflections of the child's actual behavior, was answered by comparing these reports with direct observations of the child's behavior in the home. Twenty-three children, ranging in age from 3 to 18 months, were observed at home over a two- to three-hour period. In 18 cases two separate observations were done and in 5 only one observation was done. The observation protocols were scored for temperament, using the same criteria as for the parent interviews (Thomas, et al., 1963), and the scores were compared with those obtained from the parent interview report closest in time to the observation. The comparisons showed agreement at the .01 level of confidence, permitting us to conclude that the data of the parent interviews were a valid reflection of the child's behavior.

In the collection of data on the children's behavioral characteristics, we adhered strictly to the following principles to insure the maximum validity, objectivity and reliability of the information obtained:

1. The parent and teacher interviews focused on the details of daily living during feeding, play, sleep, etc. Behavior was described in factual descriptive terms with a concern not only for what the child did but how he did it. Statements as to the presumed meaning of the child's behavior were considered unsatisfactory for primary data. When such interpretive statements were made by a parent or teacher, the interviewer pressed for an actual description. Thus, to a parental report that "the baby hated his cereal" or that "he loved his bath," the question was always posed, "What did he do specifically that made you think he loved or hated it?" Similarly, if a teacher commented that "this child always gets angry if he doesn't get his way," she was asked to give several examples with detailed descriptions of the manner in which the anger was expressed. If a staff observer reported that a child "was afraid to

ask the teacher for help," she was instructed to spell out in detail the incidents she had observed and describe the behavior she had interpreted as "fear."

2. Special emphasis was placed on recording a child's first response to a new stimulus and his subsequent reactions on re-exposure to the same stimulus until a consistent long-term response was established. Such stimuli might be simple, as the first bath or the introduction of a new food, or they might be complex, as the move to a new home, the introduction of a new person into the household or the first contact with nursery school. The sequence of responses to new stimuli, demands and situations, whether simple or complex, was found to give especially rich information on a child's individual temperamental pattern.

3. The contamination of the data collection by "halo effects" was avoided by using different staff members for different phases of the data collection for any specific child. Thus, the parent interviewer did not do the teacher interviews or direct observations, and the same staff member never did both the teacher interview and direct school observation on the same child.

4. Quality control of the interviews and observations was established by periodic checks of intra- and inter-interviewer and observer reliabilities. The item scoring of the behavior protocols also served as a continuous check on the quality and quantity of the data in each record. Interview protocol forms were revised when necessary to make them appropriate for succeeding age-periods and were pretested on samples of children not included in the longitudinal study. Quality control of the scoring procedures was also maintained by imposing periodic intra- and interscorer reliability checks.

Scoring for Temperament. Parent interviews were conducted at 3-month intervals during each child's first 18 months of life, then at 6-month intervals until five years of age, and yearly thereafter. A weighted

score model, which takes into account the item scores in all of the three scale positions used for scoring each temperamental category, was adopted. In this method, one of the extreme scale positions was represented by "0," the middle position by "1" and the other extreme position by "2." The number of items scored at each position was multiplied by the scale value (0, 1 or 2), and the products were summed. The sum of the three products was then divided by the total number of scored items in the category. Thus, a child's weighted score in any category can range from 0 to 2. This scoring method takes into account the distribution of scored items in all three scale positions, decreases the likelihood of tied scores and is both efficient and relatively simple to apply to the data.

The weights, 0-1-2, assigned to each of the ratings in the categories are, of course, arbitrary, and the question of the appropriateness of such values cannot be unequivocally resolved. This problem, however, is shared by other scoring methods, such as the per cent rank method, which also makes arbitrary assumptions as to the value of distances between rank positions. Although the issues regarding the conversion of ordinal scales to interval scales and the relative merits of parametric and nonparametric statistical analyses have received considerable attention (Cohen, 1965), there is as yet no general agreement as to an acceptable solution. We have, therefore, in the interests of simplicity and parsimony, assigned equidistant weights to the ratings in each category as the most straightforward way in which to identify polar extremes and a middle rating (the latter is defined by an intermediate term or by the occurrence in a behavioral description of equal elements of both polar extremes). To assign other than equal intervals between the ratings would in such circumstances be no less arbitrary and far more cumbersome.

Using the weighted score method, the data have been subjected to a number of quantitative analyses. For each of the first five years of life, weighted and standard scores are available for each child for each of the nine categories of temperament. Trait clusters have been identified through a principal components rotation of

three factors to a Varimax solution in each of the five years. In addition, developmental (trend) data have been computed for each child for the first five years of life. These data comprise linear and quadratic components of the scores, as well as a deviation (scatter) score.

A summary of the temperamental characteristics of our sample of children is presented in Table 2. In year 5, the mean scores of our children would describe them in terms of our categories as follows:

Activity: moderate to high moderate

Rhythmicity: regular to variable

Adaptability: adaptive to variable

Approach/Withdrawal: approaching to variable

Threshold: moderate to low moderate

Intensity: moderate to low moderate

Mood: variable to positive variable

Distractibility: distractible to variable

Persistence: persistent to variable

The means for years 1 - 5 overall tend to be similar to year 5, with the latter showing only a bit more regularity and intensity. The variability in each category is also lower in years 1 - 5 than in year 5 alone.

Task Orientation. The method used to analyze response styles to cognitive demands arose from the objective possibilities for responding and has been described in detail elsewhere (Hertzig, et al., 1968). The data utilized were the style of the child's responses to the separate test items on a standard IQ test. When confronted with a demand, a youngster is given the choice of either working or not working, and each of these may

Table 2

Means and Standard Deviations of Temperament Scores for Year 5 and Overall Years 1 through 5

Temperament Category	Temperament Scale		Year 5			Years 1 - 5		
	0	2	N	Mean	S.D.	N	Mean	S.D.
Activity	High	Low	114	0.80	.20	109	0.80	.12
Rhythmicity	Regular	Irregular	112	.48	.31	109	.60	.19
Adaptability	Adaptive	Non-Adap.	114	.59	.24	109	.63	.15
Approach/Withd.	Approach	Withdraw	114	.43	.38	109	.52	.18
Threshold	High	Low	114	1.37	.32	109	1.37	.15
Intensity	High	Low	114	1.20	.26	109	1.05	.14
Mood	Positive	Negative	114	.83	.23	109	.98	.13
Distractibility	Distract.	Non-Distr.	100	.33	.53	109	.36	.21
Persistence	Persistent	Non-Pers.	114	.58	.31	109	.50	.16

be expressed through either verbalization or action. If the choice is to work, whether verbally or nonverbally, the response may be delimited and restricted to the defined requirements of the task or may extended beyond these limits in the form of spontaneous associations and other extensions. Similarly, not-work decisions could be limited or extensive. Clearly, too, a child's initial response could either be continued or altered in response to the reiteration of demand by the examiner. Thus, initial work responses could be followed by refusals to work, verbal responses, action or any other pattern of redundancy or change. In our analyses, therefore, we classified initial responses as well as sequential response chains.

There were basically 14 categories used to define the child's responses to the demands of the Stanford-Binet tests administered at ages 3 and 6. These were:

1. Total responses expressed as Work (the child attempted to do what was asked of him).

2. Initial Work responses.

3. Initial Work responses followed by Not-Work.

4. Initial Not-Work responses followed by Work.

5. Total responses expressed verbally.

6. Work responses expressed verbally.

7. Not-Work responses expressed verbally.

8. Verbal Not-Work responses -- negation (the child says directly that he will not do what is asked of him).

9. Verbal Not-Work responses -- competence (the child says that he cannot or is unable to do what is asked of him).

10. Verbal Not-Work responses -- substitution (the child says he'll do something other than what is asked of

him; "I want to play with the toys now" when asked to draw a circle).

11. Verbal Not-Work responses -- aid (the child requests assistance).

12. Nonverbal Not-Work responses -- negation (the child gives a motor response that directly reflects refusal to participate: for example, shakes head no when asked to string beads or pushes card away when asked to tell a story about it).

13. Nonverbal Not-Work responses -- passive (the child sits still or stares ahead of him when asked to do something or presented with an activity).

14. Nonverbal Not-Work responses -- substitution (the child engages in an irrelevant physical activity rather than the one requested of him).

A child's responses to the IQ test demands were categorized according to this classification, and a percentage score was computed for each of these categories utilizing the following formula:

$$\frac{\text{Number of Responses in Category}}{\text{Number of Stanford-Binet Items}} \times 100.$$

This was modified to calculate the initial responses:

$$\frac{\text{Number of Initial Work (or Not-Work) Responses}}{\text{Total Number of Initial Responses}} \times 100.$$

All protocols from the 3 and 6 year testings were scored and percentages calculated. Intra- and interscorer reliabilities of .9 and over were established.

For the categorization of task orientation, the percentages of total work and initial work were utilized. These percentage scores indicate the degree to which a child who is presented with the request for task performance when other choices of functioning are available to him responds by addressing himself to the task.

IQ Scores. Information on each child's intellectual functioning was obtained through the administration of the Stanford-Binet, Form L, at 3 years and 6 years. In some of the children a satisfactory test performance was not obtained at 3 years. In these cases the test was repeated 6 months to a year later

At both 3 and 6 years the IQ scores of the group showed a normal distribution pattern around a mean at 3 years of 123.4, with an SD of 16.5; and at 6 years around a mean of 127.3, with an SD of 12.1. The correlations between the scores at these two test years was .64 (Hertzig and Birch, 1971).

Behavior Problem Cases. A standard clinical psychiatric evaluation was done by the research staff child psychiatrist in all cases in which suspiciously deviant behavior was reported by the parent, school or testing psychologist. On the basis of the clinical evaluation and supporting data, a judgment as to the presence or absence of a behavior problem was made. Forty-two children were identified as having significant behavior problems. In these 42 cases, data are available as to age at onset of the disorder, nature of the symptomatology, diagnosis, treatment and course on follow-up. Detailed quantitative and qualitative analyses of the relationship between temperament and environment and the ontogenesis and evaluation of the behavior problem have also been completed for all 42 children (Thomas, et al., 1968).

Other Data. Complete records of the physical health of each child, including medical and hospital reports of any serious illnesses, were obtained. A great deal of information on parental attitudes and child-care practices was also available, which enabled us to consider this important influence on a child's development. Finally, there were much miscellaneous data on age of attainment of developmental landmarks and special environmental events, as well as the results of school observations, teacher interviews and tests of cognitive skills. These data have not been utilized in the systematic quantitative analyses; however, they have proved useful in qualitative studies of those children whose academic functioning shows striking deviations from the norm.

New Data Collected

Academic Achievement. The use of standardized tests for the measurement of academic achievement is deceptively simple. As formal tests, they provide a structure that is procedurally advantageous; administration and scoring are carefully outlined and norms are based on a fairly broad sample. However, some of the shortcomings of these tests in research are too easily overlooked and their use demands caution.

To begin with, the tests are most often administered by a teacher in a classroom and this cannot be expected to provide the rigid controls that should be employed; corridor disruptions, loudspeaker announcement during the tests, as well as modifications in the timing of sub-tests are frequently found. Although the researcher using test data provided by a school system cannot control these sources of error, he certainly must not ignore them.

There is, however, a more serious problem plaguing the investigator who uses school data: he is faced with scores from both a wide range of tests and from tests that have been administered to different children at various points in their school careers.

In the bulk of research on academic achievement, this problem is circumvented. Generally, the research is limited to only one achievement test or to different batteries of the same test (e.g., Kay, 1967). If out of necessity or accident two different tests are used, the results are treated separately in statistical computation. For example, in McGhee and Crandall (1968), the Iowa Test of Basic Skills was used for grades 3, 4 and 5, while the California Achievement Test was given to grades 6, 8 and 12. In their study, the analyses of variance were done separately for each group of grades. This technique, however, is not applicable where the number of children given the same test at the same age is too small to allow quantitative analysis.

When working with these tests, too, it is absolutely necessary to avoid the error of assuming that the grade-point scores achieved in different tests are

equivalent and can be substituted for one another. Unfortunately, in routine decisions about children's school progress or school management, the grade-point average is often viewed as an absolute and valid score, regardless of the standardized achievement test employed. This is a very dangerous illusion -- in school administration and in research. The fact that two tests may be highly correlated, does not necessarily mean that their absolute scores are equivalent. A high correlation certifies only that their relative scores are consistent. For example, in Tait's study (1955) the Progressive Arithmetic and Stanford Average Arithmetic Tests were administered to a group of fifth-grade children. The two sets of tests scores yielded a correlation of .99 corrected for attenuation, yet the mean grade-point scores were 5.2 and 4.1 respectively. On the average the grade-scores reflected in the Stanford were more than a year lower than the Progressive Arithmetic Test scores. In the 8th grade Tait used two reading comprehension tests and again found almost a one year difference in the mean grade-point score; 7.7 for the Progressive Reading Comprehension Test and 6.8 for the Iowa Reading Comprehension Test, despite a correlation of .89 corrected for attenuation. Taylor and Crandall (1962) in a study of norm equivalencies among tests approved for the California State Testing Program report that "children who took the CAT (California Achievement Test) consistently received higher (grade-point) scores, while children who took the MAT (Metropolitan Achievement Test) and the SAT (Stanford Achievement Test) fairly consistently received lower grades" (1962, p. 192). For example, in 5th-grade girls with IQ's from 91-110, the mean grade for the MAT was 4.7 and for the CAT was 6.2. In Finley's study (1963) this type of result received further confirmation. Third-grade children were tested in arithmetic and reading using both the Metropolitan and California tests. The average grade-point scores on the former were lower, and the difference between the scores on the two different tests were statistically significant beyond the .01 level of confidence (Analysis of Variance).

It thus becomes quite clear that grade-point scores derived from different tests are not equivalent and cannot be used to compare children's achievement levels.

If an investigator using school records has to rely on various grade scores, his data will be subjected to the vagaries of the particular tests and the age-grades at which they were administered. Furthermore, if he only compared children who had been administered the same tests at about the same ages, most samples of children would be splintered into many small subsamples, and the impact of the results would be diminished.

However, there is a way out of this dilemma that is both practical and valid. Rather than utilize grade-point scores, we suggest percentile scores as a far better alternative. Percentile scores enable the researcher to convert all scores obtained to a common equivalent base. Different tests of academic achievement are fairly well correlated with one another (although grade scores are not equivalent) and their norms are based on similar national samples. We can, therefore, expect percentile scores to be more equivalent, since they are based on relative achievement rather than absolute scores. In addition, the percentile score yields a level of achievement measure that is reflected in a single score. The more conventional grade score has little meaning in assessing a child's performance independent of actual or expected grade placement. Thus, the percentile score achieves two very important data goals: it can be used across tests and can be integrated independently of other school data for research purposes.

This approach, whereby scores from different tests can be treated as equivalents, has still further value as scores may not only be directly compared (for example, the percentile score achieved on the Metropolitan versus that on the Stanford Achievement Test), but these scores may also be combined or averaged to add reliability to the single score that is used to characterize a child's level of performance.

The children in the present study entered first grade between 1959 and 1966 due to the cumulative collection of the sample. Although they all attended schools where standard achievement tests were administered frequently, over the span of time the schools have changed

the particular tests and the grades during which they were administered. Thus, there was much variety in the number of test scores available for each child and the ages at which they applied.

Nevertheless, a member of the research staff tried to establish personal contact with every school each child had attended to obtain his records. This required extensive effort, since many youngsters had attended a number of schools due to family moves. Some schools lacked data from previous schools the child had attended and did not know about scores in earlier grades. Others could not or would not retrieve the information for us. Out-of-town schools, which could not be visited, sometimes would not reply to our written or telephoned requests for test scores. All in all, obtaining this data was a very large task, both in terms of the numbers of children and schools involved and the variety of tests used.

As a result of the wide array of tests and range of test ages of the youngsters in our study, we were not able to generate large enough subgroups of children who had taken the same test at the same age or grade level for statistical analysis. If we had attempted to pool the children in this way we would have had too many subsamples with very few children in any one group.

To deal with this problem, we combined the data of all tests administered by the school in grades one through three for a single percentile score in arithmetic and a single percentile score in reading. In all cases, the percentile scores were based on national norms, not on the distribution of scores within our sample. For example, the school records of a particular child indicated that the Metropolitan Achievement Test was administered in Grade 1, the Metropolitan and Stanford Achievement Tests in Grade 2 and the Stanford in Grade 3. The reading scores achieved by the child on each of these tests were converted into their respective percentile scores. The sum of these four percentile scores divided by four yielded a mean percentile score for Grades 1 - 3. The same was done with his arithmetic scores. These were used to characterize the child's achievement at that point in his school career.

This procedure was also followed for Grades 4 through 6. By combining the scores for all the tests administered in the first six grades in this way, we were able to derive four academic achievement scores: Reading and Arithmetic for Grades 1 - 3 and for Grades 4 - 6.

This procedure achieved the following:

1. It avoided the problem of nonequivalent grade-point scores, since percentile scores yielded equivalent scores from different tests.

2. No data were lost, and the sample of children did not have to be fractionated by specific test or age at administration. All tests were used and all children for whom some data were available could be included.

3. The use of the average percentile score to characterize a child helped to reduce the error which may occur when a single test performance is used to characterize a child's achievement. Any one test administered on any one day by a teacher in the classroom may be of questionable reliability. However, a number of tests administered over a succession of school grades can be expected to have increased reliability.

Of course, we recognize that these are, nevertheless, performance scores and thus somewhat limited in scope and implication. But within these restrictions, the mean percentile score is an improvement over grade-point scores in research utilizing school records.

In most cases, an estimate of a child's achievement was based on more than one test score. The distribution of the number of scores available for each child is summarized in Table 3. For example, the average number of reading test scores available for the youngsters for whom we had some data from first through third grades was 3.2; for some children in this group we had as many as 7 or 8 test scores. In Grades 4 through 6, we had an average of

Table 3

Distribution of the Number of Achievement Test Scores in the School Records of Each Child, for Grades 1 - 3 and 4 - 6

Number of Test Scores	Grades 1 - 3		Grades 4 - 6	
	Reading	Arithmetic	Reading	Arithmetic
1	5	11	7	13
2	17	23	20	19
3	15	10	16	17
4	9	6	9	12
5	6	1	10	5
6	3	2	9	3
7	1	0	3	1
8	1	0	4	2
9	0	0	0	6
10	0	0	0	2
11	0	0	0	1
12	0	0	0	1
No. of Children:	57	53	78	82
Mean Tests:	3.2	2.4	3.7	3.8
Range:	1-8	1-6	1-8	1-12

3.8 arithmetic tests available for the children; in one case, there were 12 test scores available.

In addition to the fairly large range in number of test scores, there was also variety in the specific tests used. The following were those most commonly administered: Metropolitan Achievement Test; New York State (Reading and Arithmetic) Test; Iowa Test of Basic Skills; Junior Stanford Achievement Test; Stanford Achievement Test; California Achievement Test; Gates Primary Test; and New York Reading Readiness Test.

Since we also wanted an accurate estimate of the children's performance at the actual time of the study, the research staff psychologist individually administered the Wide Range Achievement Test (WRAT) to as many of the youngsters as could be seen, regardless of their grade in school. In those cases where a substantial discrepancy was found between the results on this test and the child's scores on the school-administered achievement tests, a battery of individual standard achievement tests was given to the youngster.

The general level of academic achievement, in terms of percentile scores, for these tests and grades are summarized in Tables 4 and 5. Most of the children scored above the 80th percentile in reading and arithmetic in Grades 1 - 3 and 4 - 6, and the mean achievement percentiles were between 75.3 and 83.6. In the earlier grades the mean is higher for arithmetic than reading, but in the upper grades this is reversed. However, the differences are slight and not statistically significant.

When the WRAT was administered during the course of the present study, the mean percentile in Reading was 90.4 while in Arithmetic it was 62.5. However, the variability in Arithmetic was considerable (SD = 25.1), while the variability in Reading was smaller (SD = 13.0).

Psychometric Evaluations. One of the variables that is used conventionally to characterize school children in relation to academic functioning is measured intelligence. The Wechsler Intelligence Scale for children was used in

Table 4

Distribution of Academic Achievement Scores for Reading and Arithmetic -

Scores are Average Percentiles Based on All Tests Administered in the Schools to Each Child in the Grades Indicated

Mean Percentile Score	Grades 1 - 3				Grades 4 - 6			
	Reading		Arithm.		Reading		Arithm.	
	N	%	N	%	N	%	N	%
96-99	10	17	12	23	23	28	16	19
91-95	17	28	12	23	20	25	14	17
81-90	12	20	10	19	26	32	17	20
71-80	7	12	7	13	12	15	8	10
61-70	4	7	9	17	3	4	7	8
51-60	4	7	2	4	4	5	8	10
41-50	2	3	0	0	5	6	4	5
31-40	1	2	0	0	0	0	6	7
21-30	1	2	1	2	2	2	2	2
11-20	2	3	0	0	2	2	2	2
Total:	60		53		81		84	
No Inform.:	54		61		33		30	
Mean:	80.7		83.6		81.5		75.3	
S.D.:	20.8		15.8		19.6		22.7	

Table 5

Distribution of WRAT Scores for
Reading, Spelling and Arithmetic -

Scores are Expressed as Percentiles Based on the Age of the Children at Time of Testing in 1969-70 (the Ages of the Children are Summarized in Table 1)

Percentile Score	Wide Range Achievement Test					
	Reading		Spelling		Arithmetic	
	N	%	N	%	N	%
96-99	31	55	12	22	7	13
91-95	12	21	9	16	3	5
81-90	6	11	9	16	7	13
71-80	2	4	10	18	6	11
61-70	3	5	5	9	8	15
51-60	1	2	3	5	1	2
41-50	1	2	2	4	6	11
31-40		0	3	5	12	22
21-30	0	0	0	0	4	7
11-20	0	0	2	4	0	0
-10	0	0	0	0	1	2
Total:	56		55		55	
No Data:	58		59		58	
Mean:	90.4		77.7		62.5	
S.D.:	13.0		21.9		25.1	

gathering data on the current intellectual functioning of the children. This test was chosen because of its several advantages with older children that we were anxious to exploit. The division of items into distinct areas of functioning (although it is clearly recognized that basic psychological abilities are not so precisely separated) provided an opportunity for more clear-cut analyses and comparisons of performance. In addition it permitted the determination of an IQ score that has the same interpretation at all ages, which the Stanford-Binet, Form L, does not. Although earlier testing of the children was done with the latter, the high correlation between the two scales (.82) minimized the problems inherent in the changeover.

Individual tests were administered by our own research staff psychologist. Whenever possible, testing was carried out in the same room in which the youngster had been tested at age six years. This involved a good deal of planning and time, since the children in the NYLS are now scattered throughout the suburban areas surrounding New York or even more distant. Many who had lived in the City have moved to surrounding towns and those who were originally suburbanites have frequently moved even further away. In instances where the families found it impossible to bring the child in for testing, efforts were made to see them at their homes. Trips were made to places as far away as Philadelphia so as to reduce geographical interferences with sample maintenance.

Despite our attempts, however, not all the children were seen. Although this was mostly due to the travel distances involved, in a few instances it resulted from the reluctance of these youngsters, now adolescents, to agree to testing. When they were younger, their cooperation was mandated by parental agreement to participation in the study. Now, however, as other investigators have also found, the children, themselves, must actively participate in both the decision to be studied and the study procedures. Although almost all were eventually willing to cooperate, on some occasions convenient testing sessions could not immediately be arranged and so some youngsters were not seen by the time data analysis was begun.

In our analyses of the tests administered at ages 3 and 6, we did not use an IQ score, per se, because this score on the Stanford-Binet (Form L) does not have a stable mean and standard deviation from year to year. Pinneau (1961) has demonstrated that the relative or deviation score for any given IQ score varies from year to year and has shown that a deviation score is more valid for purposes as ours. Thus, we converted all of the Stanford-Binet IQ scores to deviation scores based on the children's ages and scores. In addition, we converted these deviation scores to percentile scores so that a child's relative achievement on each test is easily ascertained. Inasmuch as the school achievement test results were also converted to percentile scores, this simplified comparisons between the two at different age-periods.

The WISC scores obtained during the current testing sessions were also converted to percentile scores, again based on national norms, not on the distribution of scores within our sample. A summary of the percentile IQ scores achieved by our study children at 3 years and 6 years on the Stanford-Binet and at a point in time between 10 and 15 years of age on the WISC is presented in Table 6. Since the WISC was not administered at a fixed age but in the calendar years 1969 - 1970, the actual ages of the children varied. This range is summarized in Table 7.

Self-Concept and Sense of Environmental Control: Achievement Motivation Questionnaire. Various methods for assessing the influence of motivation on scholastic achievement have been reported (Crandall, et al., 1962; 1965; Brown, 1966; Cohen, et al., 1959; Katz, 1967; Coleman, 1966). The factors which have been found to be most significant in these studies have been those which measured self-concept and sense of environmental control. Questionnaires have been developed for assessing these issues and were used extensively by Coleman in his report, Equality of Educational Opportunity, (1966). He and his colleagues found that there was a significant relationship between these variables and academic achievement.

Table 6

Distribution of IQ Percentile Scores Based on Pinneau (1961) and Wechsler (1951)

IQ Percentile Score	Stanford-Binet (L)				W.I.S.C. - "1969"					
	3 Years		6 Years		Verbal		Perform.		Full Sc.	
	N	%	N	%	N	%	N	%	N	%
96-99	31	31	54	51	45	57	10	12	31	39
91-95	11	11	22	21	20	25	16	20	22	28
81-90	22	22	14	13	9	11	18	22	12	15
71-80	12	12	4	4	1	1	8	10	7	9
61-70	9	9	4	4	2	3	7	9	2	3
51-60	7	7	2	2	1	1	7	9	3	4
41-50	2	2	3	3	1	1	8	10	2	3
31-40	4	4	1	1	0	0	2	2	0	0
-31	3	3	1	1	0	0	4	5	0	0
Total:	101		105		79		80		79	
Mean:	81.5		90.4		93.3		75.3		89.5	
S.D.:	18.9		15.3		9.4		21.8		12.1	

Table 7

Distribution of the Children's Ages at the Time of the 1969-70 Testing (WISC and WRAT)

Years and Months	N	%
10:0-10:5	1	1
10:6-10:11	3	4
11:0-11:5	5	6
11:6-11:11	4	5
12:0-12:5	3	4
12:6-12:11	4	5
13:0-13:5	7	9
13:6-13:11	37	47
14:0-14:5	14	18
14:6-14:11	1	1
Total:	79	

In the present study, it was not feasible to use the entire Coleman questionnaire because the subjects were also given the WISC and WRAT at the same session. (It would have been difficult, if not impossible, in most cases to have the subjects come for more than one testing appointment.) As a result, items were selected from the Coleman report by our consultant for educational research (Dr. Edmund Gordon) that were judged to be most significant for the assessment of self-concept and sense of environmental control, and these were administered to the study subjects as part of the psychological testing session. In addition, some of the questions were modified to make them more appropriate to the socio-cultural composition of our sample.

We used the following questions:

Self-Concept

1. How good a student do your teachers expect you to be?

   a. One of the best students in my class.
   b. Above the middle of the class.
   c. In the middle of my class.
   d. Just good enough to get by.
   e. Don't know.

2. I sometimes feel that I just can't learn.

   a. Agree.
   b. Not sure.
   c. Disagree.

3. I would do better in school if my teachers didn't go so fast.

   a. Agree.
   b. Not sure.
   c. Disagree.

4. I am able to do many things well.

   a. Agree.
   b. Not sure.
   c. Disagree.

Sense of Environmental Control

1. Good luck is more important than hard work for success.

   a. Agree.
   b. Not sure.
   c. Disagree.

2. Every time I try to get ahead, something or somebody stops me.

   a. Agree.
   b. Not sure.
   c. Disagree.

3. People like me don't have much of a chance to be successful in life.

   a. Agree.
   b. Not sure.
   c. Disagree.

Each subject was told that we "would like to know how you feel about work and working in school" and was asked to select an answer for each item from the choices given.

A composite score was obtained for each of the two measures by scoring each item on a three point scale (1 - 2 - 3) and then obtaining an arithmetical average for each variable. The scoring was organized so that a high average indicated a high sense of environmental control and high self-concept while a low score indicated low levels of environmental control and self-concept.

Because the study subjects were scattered through a number of school grade levels when the questionnaire was administered, we divided them into four subgroups in order to have a substantial number of subjects for analysis in each. The groups were: Boys in Grades 5, 6 and 7; Boys in Grades 8, 9 and 10; Girls in Grades 5, 6 and 7; and Girls in Grades 8, 9 and 10. Correlations were then obtained between the achievement motivation scores and achievement test scores for each group.

Methods of Data Analysis.

Quantitative Analysis. The quantitative analysis of the data made use of straightforward statistical techniques. It should be noted, however, that we did make procedural innovations in handling IQ and achievement test scores, and these have already been described.

Primarily, we did correlational analyses to see if any of the variables involved was predictive of later academic achievement. In addition, contemporaneous correlations were also calculated. While these methods were used for the sample as a whole, X^2 tests were used where applicable in examining subgroups of children in the study.

Qualitative Analysis. In this and the related NYLS studies, we have utilized quantitative methods to their maximum. Scoring methods and analytic procedures have been concerned with the fullest possible exploitation of the data by statistical techniques.

However, it is also necessary not to become so intrigued by the power and neatness of quantitative analysis as to overlook the fact that it is the product of routine methods involving a minimum of judgment and evaluation. Judgment is involved in the development of categories and the establishment of scoring criteria and methods. Once established, however, quantitative methods must of necessity be routine.

The demand for reliability in scoring also limits the possibility of the full utilization of the available data. The rigors of quantitative methods of data treatment often preclude the identification of meaningful subtleties in the performance of individual children. It therefore becomes desirable to supplement the routine quantitative methods of scoring and analysis by qualitative judgmental methods employed by professional persons with psychological and psychiatric skill and competence.

Qualitative analysis of the data was found to be especially useful in the earlier studies of temperamental

constellations and in the anterospective analyses of the child-environment process involved in the ontogenesis and evolution of behavioral disturbances. In this study, qualitative analysis was pertinent to an investigation of various influences on academic achievement in those children with behavior problems or with achievement scores at the low or high extremes of the group as a whole.

FINDINGS

Temperament and Academic Achievement

When the temperament scores in year 5 are compared with the children's academic achievement scores in Reading and Arithmetic at various points in their schooling, we see that 6 of the 9 categories of temperament are statistically significantly correlated with at least one of the seven measures of achievement (Table 8). The characteristics Approach/Withdrawal and Adaptability are most clearly related to these measures. For the former, 5 of the 7 correlations are statistically significant beyond the .05 level and for the latter, 4 of the 7 are. Thus, children who tend to be Approachers are more likely to have higher academic achievement; the correlations range from -.10 to -.38. Only the correlations with Reading and Mathematics in Grades 4 - 6 are not significantly related to the Approach/Withdrawal data. It is interesting to note, too, that the correlations were higher with the WRAT than with scores in Grades 1 - 3, although the latter were closer in time to year 5. Children who were Adaptive in year 5 also tend to have higher academic achievement.

The remaining significant correlations are widely scattered among the various measures of academic achievement: sometimes with the reading measure, sometimes with the arithmetic measure, either with measures close in time to year 5 or measures more distant in time. Without regard to relative frequency, specific skill (reading and arithmetic) or the temporal aspects, the following temperamental characteristics are related to higher academic achievement in our sample: Low Activity; Adaptability; Approach; High Threshold; Mild Intensity; Non-Persistence.

However, only with regard to Adaptability and Approach/Withdrawal can we say that the correlations are dependable: that is, are statistically significant in more than half of the measures of academic achievement in this study.

Table 8

Relationship between Temperament Scores in Year 5
and Measures of Academic Achievement

Temperament Category	High Score	Grades 1 - 3		Grades 4 - 6		W.R.A.T.		
		Read.	Math.	Read.	Math.	Read.	Spell.	Math.
Activity	Low	-.01	-.02	-.06	-.09	.26*	.12	.06
Rhythmicity	Irreg.	.01	-.03	.04	.05	-.07	-.05	-.12
Adaptability	NonAd.	-.35*	-.24	-.19	-.22*	-.26*	-.25	-.37*
Approach/With.	Withd.	-.26*	-.26*	-.10	-.18	-.35*	-.38*	-.36*
Threshold	Low	.03	-.32*	-.17	-.26*	-.11	-.04	-.10
Intensity	Mild	.28*	.20	.18	.07	.22	.28*	.22
Mood	Negat.	-.23	-.22	-.10	-.10	-.21	-.06	-.09
Distractib.	NonD.	.08	.05	.00	.05	-.22	-.08	-.03
Persistence	NonP.	.24	.27	.24*	.10	.10	.18	.05
N(Range)		51-59	47-55	73-84	73-84	51-57	50-56	50-56

*Statistically significant beyond the .05 level of confidence for the particular number of cases involved.

In the case of the remaining three temperamental characteristics (Rhythmicity; Distractibility; and Mood), the correlations are quite low. For example, the correlations between Rhythmicity and the seven measures of academic achievement range from -.12 to .05. Thus, the children's patterns of Regularity-Irregularity are not statistically related to measured reading or mathematics achievement in Grades 1 - 3, 4 - 6 or in the Wide Range Achievement Test administered as part of this study. From these data we can also see that neither the subject matter, reading or arithmetic, nor the degree of remoteness from year 5 plays a role in the relationship; Grades 1 - 3 generally are completed when the children are about 9 years old (4 years after the temperament data were collected); Grades 4 - 6 are completed when the children are about 11 to 12 years of age; and the WRAT was administered when at least half of the children were at least 12 years old.

In addition to Rhythmicity, there are two other categories that are not statistically significantly related to any of the measures of academic achievement, Distractibility and Mood. However, there is a slight but not statistically significant tendency for the children who show more positive mood in year 5 to be more likely to have higher achievement scores in Grades 1 - 3; the correlations involved are -.23 and -.22.

From these data it can be seen that the characteristics of Non-Adaptability and Withdrawal, which typify the slow to warm up child, show the most significant correlations with academic achievement. Temperamentally, this type of child retreats from new experiences or demands with either quiet withdrawal or mild complaints and shows gradual adaptation only after repeated contact with the stimulus allows him to become familiar with it. The finding that, as a group, children with these characteristics do less well on the measures of academic achievement may be due to one or more of the following explanations.

1. Their slow involvement may interfere with their adaptation to the classroom situation and to academic demands, thereby reducing their responsiveness to learning material. The slow to warm up

child typically is stimulated less and learns less in the initial phases of any new school demand, and this may be reflected in lower achievement scores.

2. Teachers may misjudge such children as having less intellectual capacity than they do, in fact, possess. As a result, they may expect fewer accomplishments from such youngsters. When teacher expectations are transmitted to the child, they may influence his actual learning so that it is below what he really can do (Gordon and Thomas, 1967).

3. The temperamental characteristics of Withdrawal and Slow Adaptability may interfere with the child's attaining a maximum level of performance on the achievement test itself. This would pertain particularly to school-administered tests. However, since the youngsters who were slow to warm up also got lower scores in the WRAT administered individually by our own staff psychologist who knew the youngsters well and who structured the test situation so as to give each child whatever time was necessary for him to adapt to the test demands, this explanation would seem to be insufficient.

The decision as to which of the above best explains the lower academic achievement of the slow to warm up children as a group cannot be made on the basis of data from the present study. Nevertheless, the above possibilities should be considered: it may be that one factor operates in some cases, another in different instances or some combination may be implicated. In any event, the findings do highlight the importance of identifying children with these temperamental characteristics in the school setting so that the special efforts required if they are to function academically at an adequate level can be made.

The remaining data provide only minimal support for a relationship between the other temperamental traits and the measures of academic achievement we used. Although a number of mitigating factors are involved, their impact can only be assessed indirectly. For example,

the measures of temperament are being employed in a predictive test, and any simultaneous relationship is not measured. In addition, the temperamental characterizations are based on data culled from parental reports and describe the children at home interacting with their parents and siblings rather than with their teachers and classmates in the school situation. There is much to suggest that the latter may be the more important feature when home and school patterns are not identical (Gordon and Thomas, 1967). Furthermore, our children are homogeneous in ability. They are all above average in measured intelligence and their academic achievement is also generally high. A more adequate test, a more complete test, of the temperament-academic achievement relationship should include a wider range of ability levels. Finally, the small number of significant correlations for the study population as a whole between temperamental characteristics in the preschool years and later academic achievement does not preclude the possibility that there may be a significant relationship between them in individual children. This possibility is considered below in the discussion of the qualitative analysis of the data.

Temperament and IQ

The relationship between temperament and IQ is summarized in Table 9A. It is interesting to note that although temperament at year 5 is not statistically significantly related to the Stanford-Binet scores at years 3 or 6 (except for the Activity-IQ correlation at year 3), 5 of the 9 temperamental characteristics are significantly related to the WISC Verbal Scale score. Furthermore, 4 are significantly related to the WISC Full Scale score. Only one of the characteristics (Approach/Withdrawal) is related to the Performance Scale score.

As this table makes clear, children who are Approachers are more likely to do better on the Verbal, Performance and Full Scale WISC. This temperamental trait is also correlated with many of the academic achievement scores (Table 8). In addition, Adaptability, High Threshold and Mild Intensity are positively correlated with the Verbal and Full Scale WISC scores, while Positive Modd is positively correlated with the Verbal Scale alone. The

Table 9

Relationships among IQ Test Scores and Measures of Temperament (Year 5) and Academic Achievement

A. Temperament Category (Yr 5)	High Score	Stanf-Binet(L) Yr 3	Stanf-Binet(L) Yr 6	W.I.S.C. Verbal	W.I.S.C. Perf.	W.I.S.C. Full S.
Activity	Low	-.19*	.08	.00	-.09	-.09
Rhythmicity	Irreg.	.12	.12	-.15	.09	-.04
Adaptability	NonA.	-.15	-.19	-.28*	-.10	-.21*
Approach/W.	With.	-.13	-.09	-.30*	-.34*	-.40*
Threshold	Low	.10	-.13	-.26*	-.13	-.21*
Intensity	Mild	-.12	.18	.26*	.19	.25*
Mood	Negat.	-.10	-.11	-.20*	-.12	-.16
Distractib.	NonD.	-.07	-.01	-.18	-.15	-.14
Persistence	NonP.	-.09	.06	.13	.03	.12
	N(Range)	92-106	97-110	73-82		
B. Grades 1 - 3						
Reading		.35*	.50*	.33*	.14	.32*
Arithmetic		.16	.39*	.32*	.00	.17
	N(Range)	50-54	53-57	45-48		
C. Grades 4 - 6						
Reading		.54*	.65*	.62*	.38*	.62*
Arithmetic		.48*	.62*	.58*	.32*	.57*
	N(Range)	81	81	64-66		
D. W.R.A.T.						
Reading		.19	.36*	.62*	.34*	.55*
Spelling		.07	.14	.36*	.22	.22
Mathematics		.26	.40*	.55*	.38*	.49*
	N(Range)	52-53	52-53	56-57		

*Correlation is significant beyond the .05 level of confidence for the number of cases involved.

remaining temperamental correlations (Activity, Rhythmicity, Distractibility and Persistence) were not statistically correlated with IQ as measured on the WISC.

IQ and Academic Achievement

The correlations between the children's measured intelligence and their academic achievement scores are presented in Table 9, parts B, C and D. Of this broad series of correlations, some have a very long predictive span (e.g., 3 year IQ and WRAT), some have a shorter span (e.g., 6 year IQ and Grades 4 - 6 achievement) and some are contemporaneous (e.g., WISC and WRAT).

The IQ scores at year 3 provide the earliest base for "predicting" academic achievement. It is interesting to note that the correlations between year 3 IQ and academic achievement scores in Grades 4 - 6 are higher (.54 with Reading and .48 with Arithmetic) than those between year 3 IQ and the achievement scores in Grades 1 - 3 (.35 in Reading and .16 in Arithmetic), despite the fact that the earlier grades are 3 to 6 years from the date of the IQ test while Grades 4 - 6 are 6 to 9 years after the year 3 testing. The explanation of this "sleeper effect" is not clear from our data, but may be related to the nature of early and later school achievement demands and the adjustments required.

The IQ data obtained when the children were six years old show fairly consistent and high correlations with school-measured academic achievement. In Grades 1 - 3, the correlations are .50 and .39 in Reading and Arithmetic, respectively; in Grades 4 - 6, the correlations are .65 and .62. Obviously, the IQ achieved in year 6 is a fairly good "predictor" of academic achievement in the primary grades for this group of children.

It is worth noting that the correlations of Stanford-Binet IQ with school-derived data are higher than those found with the Wide Range Achievement Test administered as part of this study. This may be a function of the longer interval involved, since the WRAT was administered when about 75 per cent of the children had

already completed the 8th grade, or it may be due to the very limited nature of the WRAT.

The WISC data presented in Table 9 cannot be viewed as part of the "predictive" correlations with the academic achievement scores in Grades 1 - 3 or 4 - 6, since they follow rather than precede the administration of those tests. Nevertheless, the correlations between academic achievement in Grades 4 - 6 and the Verbal and Full Scale WISC IQ scores are quite high (.62, .62, .58 and .57) and those with the Performance Scale IQ are also significant, though lower (.38 and .32).

Signs of the Difficult Child and Academic Achievement

The temperamental characteristics associated with the "difficult" child include Irregularity, Non-Adaptability, Withdrawal, Negative Mood and High Intensity. This temperamental constellation is of special interest because of the high frequency with which it is associated with the development of behavior disorders (Thomas, et al., 1968; Chess and Korn, 1970). We were concerned, therefore, to determine whether there was any relationship between this temperamental constellation and academic achievement.

Those youngsters who have more signs of the difficult child in year 5 tend to do more poorly on measures of academic achievement in Grades 1 - 3 (correlations of -.20 with Reading and -.22 with Arithmetic) and they also tend to do more poorly in the later years as well (Table 10). On the WRAT there are correlations of -.16, -.14 and -.12 in Reading, Spelling and Mathematics. However, none of these correlations are statistically significant beyond the .05 level of confidence and so their interpretative significance is blunted.

This lack of statistically significant correlation may be due to factors in the children's backgrounds and environment within and outside school that soften the impact of these temperamental characteristics. In other words, there may be sufficient compensatory factors that override the negative consequences of these aspects of temperament. For example, the children are, on the

Table 10

Frequency Distribution of Signs
of the Difficult Child (Year 5)

Number of Signs in Year 5	Frequency N	Frequency %	
0	8	7.3	
1	17	15.5	Mean = 2.54 signs
2	27	24.5	SD = 1.29
3	32	29.1	
4	20	18.2	
5	6	5.5	
Total	110		

Correlations between Number of Signs
of the Difficult Child in Year 5
and Measures of Academic Achievement

Measures	N	Correlation
Reading Grades 1 - 3	56	-.20
Arithmetic Grades 1 - 3	52	-.22
Reading Grades 4 - 6	81	.02
Arithmetic Grades 4 - 6	81	.02
Reading WRAT	55	-.16
Spelling WRAT	54	-.14
Mathematics WRAT	54	-.12

average, very bright, come from educationally stimulating homes and their temperamental characteristics are, for the most part, not at the extreme ranges; the deviations we refer to, the signs of the difficult child, are based on relative scores within the group studied, not absolute scores. Thus, it may well be possible that significant correlations may be found in difficult children who are not as bright intellectually as those we studied and do not come from educationally stimulating homes.

Task Orientation, Academic Achievement and IQ

The data on the children's responses to the cognitive demands of the Stanford-Binet IQ administered when they were 3 and 6 years of age are summarized in Table 11. The 14 types of response recorded at both testing times are tabulated. (The definitions of these categories have been given earlier.)

In all of the response categories except Item 4 (Initial Not-Work responses followed by Work) the differences in the means for the two test ages were significant beyond the .05 level of confidence. As can be seen, there is a significant increase in both the work and verbal responses of the children as they got older, along with a decrease in the number of times they said they would not do something (Item 8). Although there appears to be a large increase in the frequency with which the 6-year-olds gave negation or passive nonverbal Non-Work responses, it should be recalled that these percentages are based on only a small total number of nonverbal Not-Work responses and thus the figures are relative, not absolute.

Of special interest to this study was the pertinence of the children's response styles to a cognitive demand to their academic achievement. These data are summarized in Table 12 as correlations between each of the scores in the "demands analysis" (task orientation) and the various measures of academic achievement in Grades 1 - 3, 4 - 6 and on the Wide Range Achievement Test administered as part of this study.

Table 11

Response to Cognitive Demands; 3 Year and 6 Year IQ Testing
(N = 104)

Proportion:	Year 3		Year 6	
	Mean	S.D.	Mean	S.D.
1. Total responses expressed as Work	69.4	18.3	75.0	10.2
2. Demands responded to by Initial Work	78.4	19.1	85.6	9.4
3. Initial Work responses followed by Not-Work	25.2	15.9	19.7	11.0
4. Initial Not-Work responses followed by Work	55.3	36.7	59.1	33.9
5. Total responses expressed Verbally	55.7	14.1	72.7	11.7
6. Work responses expressed Verbally	53.6	16.0	75.5	8.7
7. Not-Work responses expressed Verbally	57.8	26.8	68.7	24.3
8. Verbal Not-Work responses - Negation	22.3	24.7	12.3	15.9
9. Verbal Not-Work responses - Competence	35.7	28.9	51.8	25.1
10. Verbal Not-Work responses - Substitution	24.5	19.6	19.7	16.2
11. Verbal Not-Work responses - Aid	10.7	17.9	15.3	14.8
12. Nonverbal Not-Work responses - Negation	14.8	20.5	37.4	29.1
13. Nonverbal Not-Work responses - Passive	33.5	31.1	46.3	30.2
14. Nonverbal Not-Work responses - Substitution	45.6	31.5	10.8	20.0

Table 12

Demands: Academic Achievement

Demands 6 Yr	Grades 1-3		Grades 4-6		W.R.A.T.			S-B(L)		W.I.S.C.		
	Read	Math	Read	Math	Read	Spell	Math	3 Yr	6 Yr	V.	P.	F.S.
1	-.13	-.02	-.09	-.07	-.08	-.06	-.07	.06	.04	.03	.03	.06
2	-.23	-.26	.04	.08	-.24	-.06	-.12	.02	.02	-.03	.06	.05
3	.03	-.09	-.04	-.08	.04	.22	-.07	-.05	-.12	.08	-.21	-.06
4	.01	.08	.13	.16	-.05	-.21	.05	.08	.28*	.18	-.21	-.05
5	.07	-.08	.06	.02	-.19	-.22	-.16	.24*	.27*	-.13	-.12	-.15
6	.28*	-.04	.17	.02	-.04	.06	.09	.23*	.29*	.01	.00	.02
7	-.28*	-.07	-.07	-.06	-.15	-.27*	-.29*	.10	.09	-.13	-.22*	-.19
8	.03	.01	.04	.12	.15	.18	.17	.16	.11	.14	.01	.15
9	.13	-.12	.10	.07	-.16	-.22	-.21	.23*	-.02	-.26*	-.12	-.23*
10	-.06	.16	-.01	-.04	.11	.10	.04	-.07	.05	.19	.16	.20
11	-.16	-.08	-.24*	-.25*	-.03	.07	.07	-.03	-.13	.08	-.06	.02
12	.02	.16	.08	-.02	.03	.08	-.06	.12	.04	-.10	.09	-.02
13	.07	-.16	-.12	-.02	.04	.10	.21	-.05	-.11	.05	.11	.12
14	-.15	-.10	-.12	.00	-.14	-.25	-.11	-.15	.01	.03	-.13	-.02
N	53	49	77	77	52	51	51	98	103	77	77	77

*Significant beyond .05 level of confidence.

(continued)

Table 12 Continued

Demands: Academic Achievement

Demands 3 Yr	Grades 1-3		Grades 4-6		W.R.A.T.			S-B(L)		W.I.S.C.		
	Read	Math	Read	Math	Read	Spell	Math	3 Yr	6 Yr	V.	P.	F.S.
1	.08	.12	-.03	.03	-.02	-.04	.20	.14	.06	.17	.24*	.27*
2	.13	.12	-.03	.03	.06	-.05	.19	.10	.01	.20	.22*	.26*
3	.10	-.07	-.04	-.14	.20	.07	.00	-.17	-.10	.13	-.25*	-.05
4	-.09	.15	.05	.09	.14	.01	.14	-.02	.13	.17	.20	.23*
5	.26	.12	.21	.08	.30*	.04	.04	.13	.26*	.29*	.26*	.36*
6	.11	-.17	.25*	.09	.19	-.09	-.17	.15	.14	.17	.12	.23*
7	.18	.13	-.02	-.09	.02	-.07	.00	.05	.13	-.02	.13	.07
8	-.07	.02	.00	.06	-.04	.15	.10	-.20*	-.11	-.05	-.15	-.15
9	.20	.15	.10	.04	.17	.09	.10	.24*	.27*	.08	.17	.17
10	.06	.05	-.23*	-.20	.08	-.13	-.19	-.14	-.06	-.18	-.04	-.15
11	-.09	-.17	.14	.11	-.01	-.20	-.09	-.13	-.09	.07	-.04	.05
12	.16	-.10	.23*	.14	.02	.03	-.03	.10	.14	.13	.05	.13
13	-.05	.12	-.04	.01	.09	.19	.13	-.14	-.15	.01	-.03	-.04
14	-.19	-.23	-.16	-.19	-.18	-.17	-.26	-.11	-.04	-.12	-.12	-.13
N	53	49	77	77	52	51	51	98	103	77	77	77

*Significant beyond .05 level of confidence.

Because the matrices of correlations are rather large, 98 correlations in each, the presence of 4 or 6 statistically significant correlations in each matrix may very well be within chance expectancy when the .05 level of confidence is used. However, in examining the data for the "demands analysis" derived from testing at year 6 we find that all of the significant correlations involve only three response categories.

It appears that the relative number of "Not-Work responses expressed Verbally" (Item 7) is negatively correlated with Reading in Grades 1 - 3 and with Spelling and Mathematics on the WRAT. The latter are especially noteworthy because they involve performances on a test administered in about 75 per cent of the cases at least 7½ years after the year 6 "demands analysis" data were collected.

In addition, the relative number of "Verbal Not-Work responses invoking Aid" (Item 11) in year 6 was also inversely related to achievement measures, specifically Reading and Arithmetic scores in Grades 4 - 6 -- events occurring 4 to 6 years after the demans data were collected.

The significance of the data concerning the children's responses to cognitive demands is also seen in their correlations with IQ test scores on the Stanford-Binet at years 3 and 6 and the WISC administered as part of this investigation (Table 11).

It is quite striking to find that there are fewer statistically significant correlations between the responses to cognitive demands at year 3 and the year 3 or year 6 IQ test results than with the WISC scores at the older age-period of the present study. Only three categories were significantly correlated with the 3 and/or 6 year IQ's: Verbal Not-Work responses - Competence (Item 9) positively and significantly correlated with the year 3 and year 6 IQ data; Total responses expressed Verbally (Item 5) was positively related to year 6 IQ test scores; and Verbal Not-Work responses - Negation (Item 8) was negatively and significantly related to year 3 IQ test scores.

By contrast, six different categories of response to cognitive demands were significantly correlated with one or more scores on the WISC although this test was administered to most of the children about nine years after the year 3 "demands data" were observed. The following five categories were significantly and positively correlated with the Full Scale WISC IQ scores of the children: Total responses expressed as Work; Initial Work responses; Initial Not-Work responses followed by Work; Total responses expressed Verbally; and Work responses expressed Verbally.

All in all, the year 3 demands data yielded ten significant correlations with one or more of the WISC scores and these ten involved only the first six categories enumerated on Table 12 (and spelled out on Table 11).

The fact that these data involve a fairly long-term predictive analysis makes them quite impressive, especially since IQ stability over the same time span is relatively low. Moreover, the fact that the year 6 demands data yield only three significant correlations with the WISC (and these involve only two categories of response, neither of which correlated with the year 3 data) seems to suggest that the children's responsiveness to cognitive demands at an earlier age (three years) is more relevant to performance in middle childhood than is the cognitive response style at a later age (six years). In explanation, it may be that the year 6 data already reflect the children's adaptability to social expectancy rather than their more basic characteristics which are evident when they are younger. This may also account for the absence of significant correlations between the year 3 and year 6 demands data. The correlations for the 14 response categories ranged from -.13 to +.20, with eight correlations between plus or minus .10.

Achievement Motivation and Academic Achievement

For these analyses, the children were subdivided by grade as well as by sex because it was felt that the factors underlying the younger children's attitudes would

be different from those for the adolescents in the older groups. The analyses were done separately for sense of environment control and self concept.

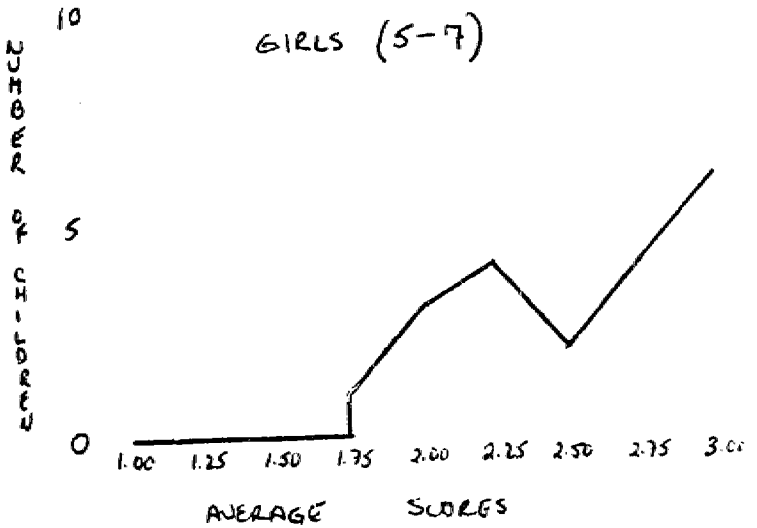
The distribution of the children's achievement motivation scores is shown in Tables 13, 14 and 15. As can be seen from these graphs, the distribution of scores is markedly skewed in all the subgroups for both measures, especially so for sense of environmental control. In all instances, the skewing is in the same direction with the heavy loading at the upper (positive) end of the scale. This marked homogeneity and lack of differentiation within the groups may stem from two factors. First, it may reflect the actual life experiences of this population who come from advantaged homes, have, with few exceptions, above average or superior IQ's, have been encouraged in their intellectual functioning at home and have attended good schools. Second, it may be that the range of questions used was insufficient to delineate differences among the children. While this second possibility could only be eliminated by re-testing with an extensive battery of questions, it seems unlikely that this by itself would account for the degree of homogeneity in the responses of the sample. This finding reaffirms the oft-repeated warning that norms established for one group should not be ipso facto considered valid for other groups.

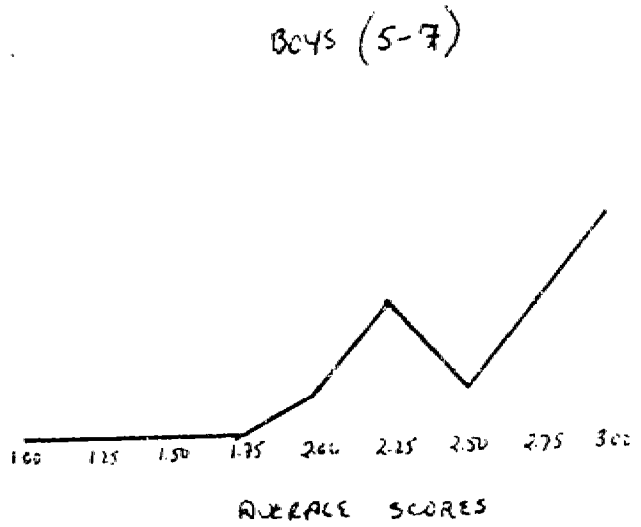
The data dealing with the children's mean self concept and environment control scores and their correlations with academic achievement (WRAT, reading and arithmetic) are summarized in Tables 16 and 17. The WRAT was used for the evaluation of academic achievement because it was obtained at the same time that achievement motivation was measured.

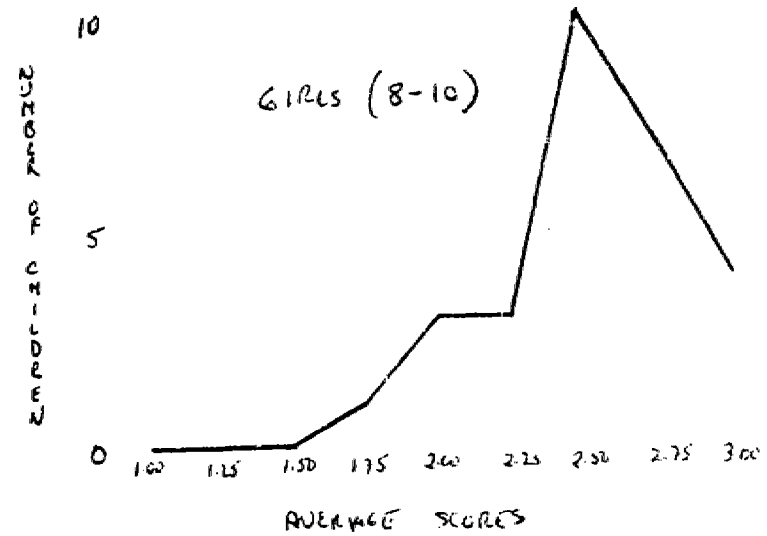
The data in Table 16 again indicate the homogeneity of the sample with regard to achievement motivation scores. The average self concept varied from 2.5 to 2.7, with 3.0 being the highest possible score that could be achieved. There are slight variations in the distributions of scores with the boys' data somewhat more skewed. But none of the children tend to score at the low extreme range. The data on control of environment are even more striking.

TABLE 13

FREQUENCY DISTRIBUTION OF AVERAGE SELF CONCEPT SCORES







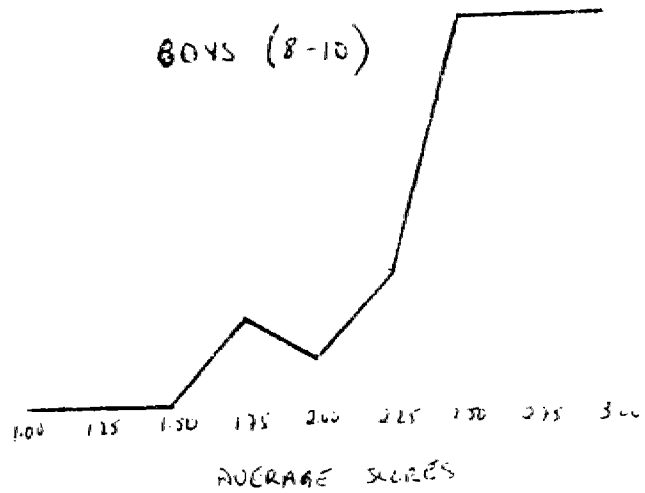


TABLE 14

FREQUENCY DISTRIBUTION OF AVERAGE ENVIRONMENT CONTROL SCORES (GRADES 5-7)

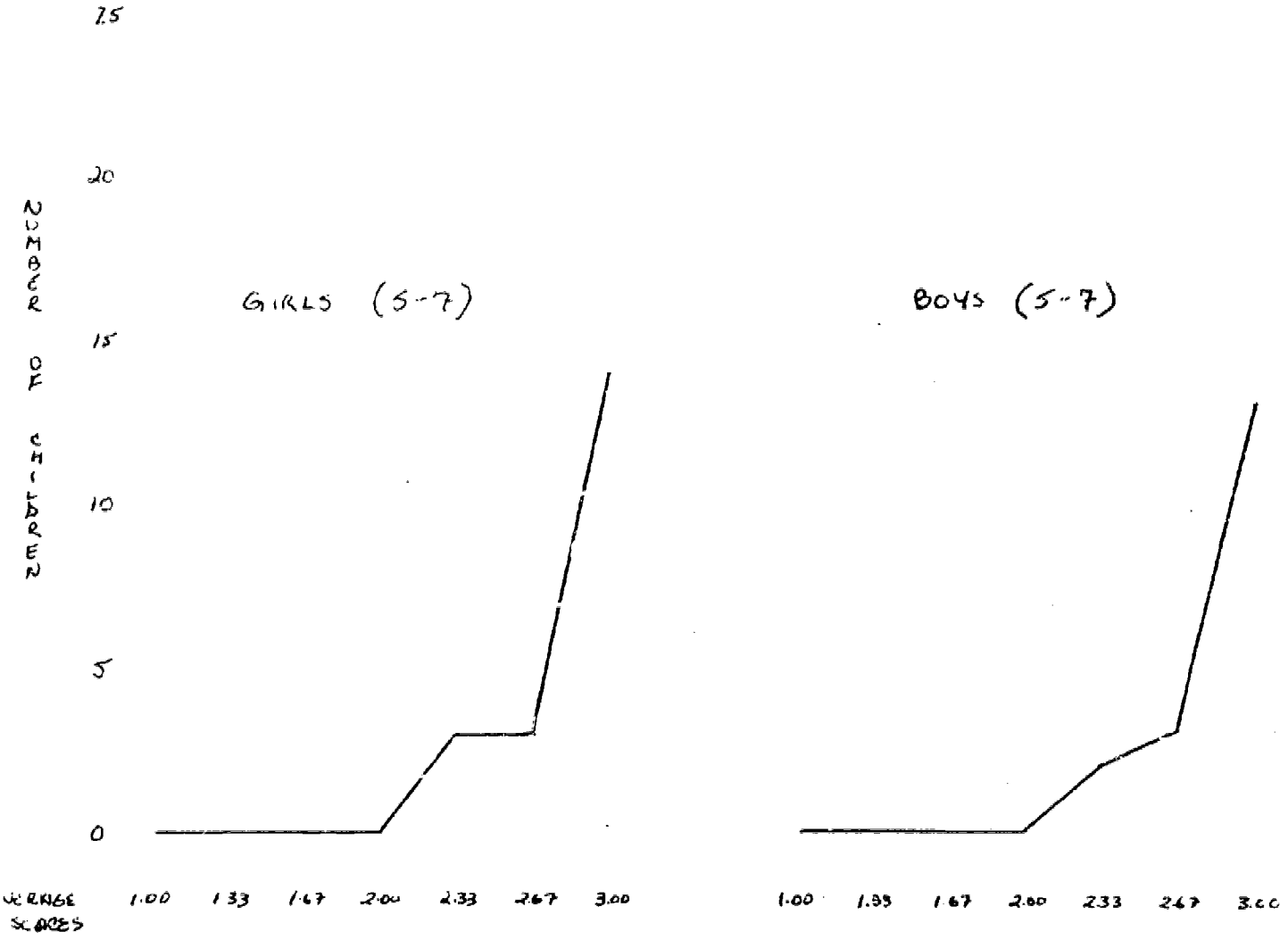


TABLE 15

FREQUENCY DISTRIBUTION OF AVERAGE ENVIRONMENT CONTROL SCORES (GRADES 8-10)

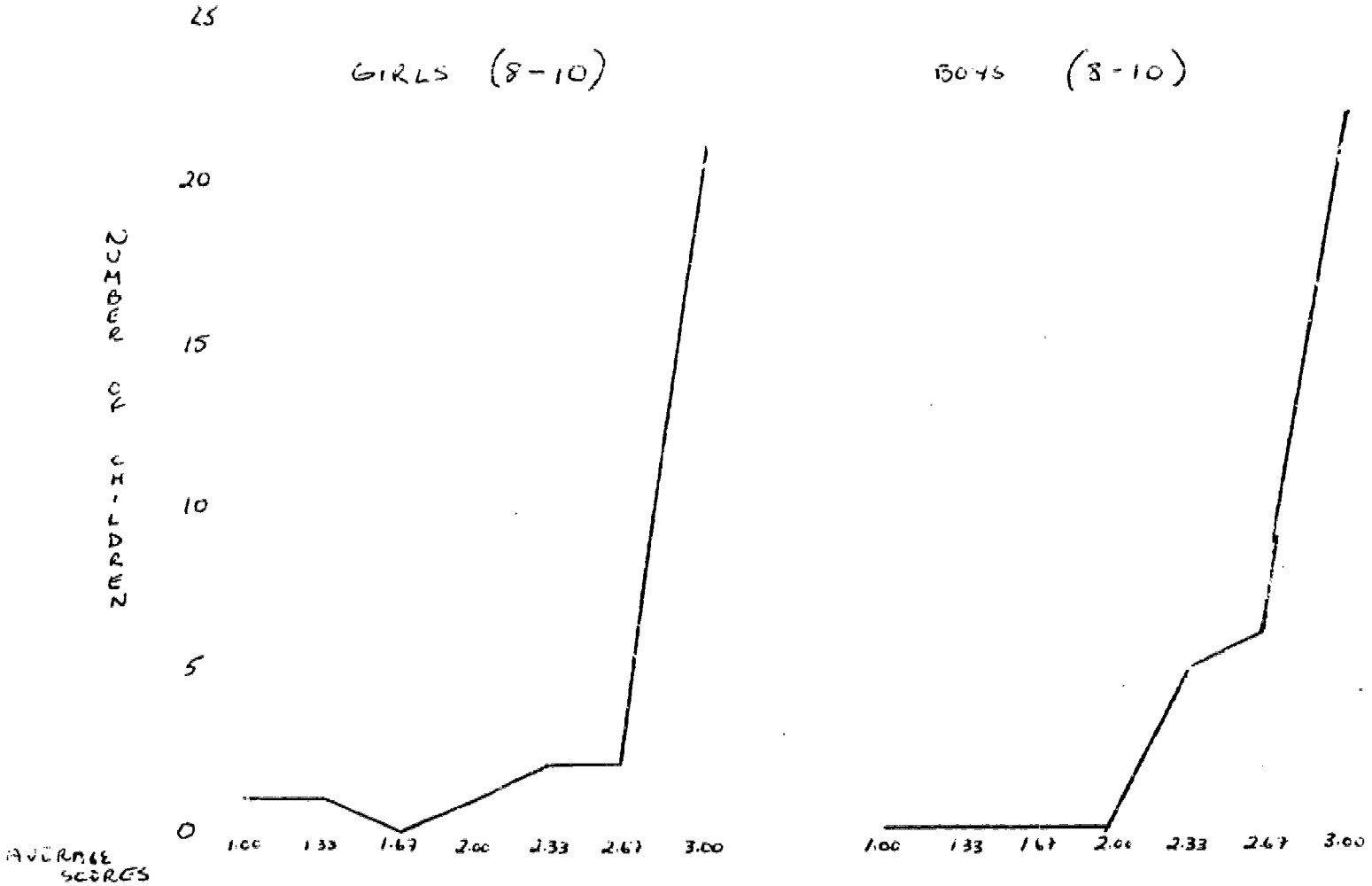


Table 16

Mean Achievement Motivation Scores and
Academic Achievement Percentiles (WRAT)

	Grades 5-7		Grades 8-10	
	Boys	Girls	Boys	Girls
Self Concept	2.6	2.5		2.5
Control of Environment	2.9	2.8	2.9	2.7
Reading (WRAT)	83.0	91.6	93.2	91.9
Arithmetic (WRAT)	58.2	56.2	70.1	60.7
Number of Cases	10	19	18	24

The range of scores is 2.7 to 2.9, again with 3.0 being the highest possible positive score.

The data on academic achievement show the children to be higher in reading than in arithmetic, with the group as a whole achieving above the mean percentile for their grade placement.

The correlations between the achievement motivation scores and academic achievement are presented in Table 17. It was felt useful to do these correlations despite the marked skewing of the achievement motivation scores. Self concept was significantly correlated with reading achievement only in the older girls; the correlation of .55 that was obtained is statistically significant beyond the .05 level of confidence. In none of the other groups were self concept and academic achievement significantly correlated.

Control of environment was significantly correlated with both reading and arithmetic in the older girls -- .44 and .39 respectively. In the younger boys' group, the correlations between these measures approach statistical significance; the .53 and .55 measures just miss the .05 level of confidence because of the small number of cases involved. The other correlations are not statistically significant.

These low order levels of correlations probably result, at least in part, from the marked skewing in the distribution of achievement motivation scores. Problems of reliability and validity may also have played a part in these results.

The children with low scores on self concept and environment control (below 2.0) were identified. For environmental control, there were two such children, and both were also in the lower third of the total population in terms of their academic achievement. One youngster was the most seriously disturbed case in the sample psychiatrically and had required several periods of residential psychiatric treatment. The other was a

Table 17

Correlations between
Mean Achievement Motivation Scores
and Academic Achievement Scores

	Grades 5-7		Grades 8-10	
	Boys	Girls	Boys	Girls
Self Concept and Reading	.12	.24	.00	.55**
Self Concept and Arithmetic	.13	.02	.11	.14
Control of Environment and Reading	.53	.19	.29	.44*
Control of Environment and Arithmetic	.55	.02	.33	.39*

*Statistically significant beyond the .05 level of confidence.

**Statistically significant beyond the .01 level of confidence.

child with marked difficulties in orienting herself to task performance of any kind, which appeared to be the result of parental handling and attitudes. Neither of these two children were in the low self-concept group.

There were five children with low self-concept scores (below 2.0). No obvious common characteristics distinguished them from other youngsters in the sample.

Qualitative Analysis

As indicated earlier, the quantitative correlational analysis of the behavior problem group was not feasible because the subsamples established in terms of age, sex, and so on were too small. Detailed qualitative culling of the records was done as an alternative procedure, as has been done in our previous studies of these children with behavior problems (Thomas, et al., 1968).

Qualitative analysis of the behavior problem cases showed that they were distributed among the high, average and low achievers. The low achievers, with few exceptions, came from homes in which there was marked intrafamilial stress, either in terms of serious marital discord leading to separation or confused or disorganized functioning on the part of the mother. One exception was a stable family in which both children were low achievers. In this case, the parents were extremely permissive in their approach. The high achievers among the behavior problem cases were mostly from stable families, though here, too, there were a few exceptions in which marital discord played a role.

Outside of this issue of intrafamilial stress, the qualitative analysis of the behavior problem cases showed no other even suggestive correlations with academic achievement, such as with type and severity of problem or age at onset of problem.

CONCLUSIONS

The major objectives of the study involved the determination of the predictive value of characteristics of behavioral style, specifically temperament and task orientation in the preschool period, for academic achievement in elementary school.

The most significant correlations found in this regard were for the temperamental characteristics of non-adaptability and withdrawal responses to new situations. These characteristics typify the child whom we have called "Slow to Warm Up." Temperamentally, this type of child retreats from new experiences or demands with either quiet withdrawal or mild complaints and shows gradual adaptation only after repeated contact with the stimulus allows him to become familiar with it. The finding that, as a group, children with these characteristics do less well on the measures of academic achievement may be due to one or more of the following explanations.

1. Their slow involvement may interfere with their adaptation to the classroom situation and to academic demands, thereby reducing their responsiveness to learning material. The slow to warm up child typically is stimulated less and learns less in the initial phases of any new school demand, and this may be reflected in lower achievement scores.

2. Teachers may misjudge such children as having less intellectual capacity than they do, in fact, possess. As a result, they may expect fewer accomplishments from such youngsters. When teacher expectations are transmitted to the child, they may influence his actual learning so that it is below what he really can do (Gordon and Thomas, 1967).

3. The temperamental characteristics of Withdrawal and Slow Adaptability may interfere with the child's attaining a maximum level of performance on the achievement test itself. This would pertain particularly to school-administered tests. However, since

the youngsters who were slow to warm up also got lower scores in the WRAT administered individually by our own staff psychologist who knew the youngsters well and who structured the test situation so as to give each child whatever time was necessary for him to adapt to the test demands, this explanation would seem to be insufficient.

The decision as to which of the above best explains the lower academic achievement of the slow to warm up children as a group cannot be made on the basis of data from the present study. Nevertheless, the above possibilities should be considered: it may be that one factor operates in some cases, another in different instances or some combination may be implicated. In any event, the findings do highlight the importance of identifying children with these temperamental characteristics in the school setting so that the special efforts required if they are to function academically at an adequate level can be made.

Interestingly, the temperamental characteristics associated with the "difficult" child did not show a relationship to academic achievement. This temperamental constellation of Irregularity, Non-Adaptability, Withdrawal, Negative Mood and High Intensity is frequently associated with the development of behavior disorders. Since subdividing the behavior problem cases into groups according to age resulted in samples too small for quantitative analysis, we were particularly concerned with determining if there was a relationship between risk of development of behavior problems (in terms of temperament) in the difficult child and risk of academic underachievement.

In this sample of children, there was no such relationship. This lack of statistically significant correlation may be due to factors in the children's backgrounds and environment within and outside school that soften the impact of these temperamental characteristics. In other words, there may be sufficient compensatory factors that override the negative consequences of these aspects of temperament. For example, the children are, on the average, very bright, come from educationally stimulating homes and

their temperamental characteristics are, for the most part, not at the extreme ranges; the deviations we refer to, the signs of the difficult child, are based on relative scores within the group studied, not absolute scores. Thus, it may well be possible that significant correlations may be found in difficult children who are not as bright intellectually as those we studied and do not come from educationally stimulating homes.

It is also of interest that there was no significant correlation found between the temperamental characteristic of distractibility at age 5 and later academic achievement and only a single correlation between one of the measures of achievement and activity level. Children with these characteristics are frequently thought of as constituting school problems and, indeed, do have difficulties in coping with learning demands when the school situation is not structured to allow for the behavioral manifestations of either high motor activity or distractibility. Again, it would appear that bright children with high academic motivation who attend good schools are capable of satisfactory academic achievement despite these behavioral liabilities.

The overall relationship between temperament and IQ was, in general, not clear-cut. While temperament at year 5 was not related to IQ scores at ages 3 or 6 (except for one correlation with activity level), five of the nine characteristics were significantly related to the WISC Verbal Scale and, of these, four were significantly related to the WISC Full Scale scores. These were Approach, Adaptability, High Threshold and Mild Intensity. While these findings again point out the vulnerability of the slow to warm up child in intellectual activities, the generally low level of significant relationships may result from the fact that the intellectual functioning of the study youngsters is basically superior, thus blunting the relevance of behavioral style.

As one would expect, correlations between IQ and academic achievement were generally high. Of special note was the fact that year 3 IQ scores were more highly correlated with academic achievement scores in Grades 4 - 6 than with

achievement scores in Grades 1 - 3, although the latter were temporally closer. The explanation of this "sleeper effect" is not clear from our data, but may be related to the nature of early and later school achievement demands and the adjustments required.

The IQ data obtained when the children were six years old show fairly consistent and high correlations with school-measured academic achievement. It was obvious that the IQ achieved at this age was a fairly good "predictor" of academic achievement in the primary grades for this group of children.

It is worth noting that the correlations of the 3 and 6 year Stanford-Binet IQ's with school-derived data were higher than those found with the Wide Range Achievement Test administered as part of this study. This may be a function of the longer interval involved, since the WRAT was administered when about 75 per cent of the children had already completed the 8th grade, or it may be due to the very limited nature of the WRAT.

In analyzing the pertinence of the children's response styles to cognitive demands at age 6 to their academic achievement, we found that all the significant correlations involved only 3 response categories. Children who gave verbal not-work responses, and particularly those who requested aid, did, on the average, less well on a variety of measures of academic achievement than the other youngsters.

The significance of the data concerning the children's responses to cognitive demands was also seen in their correlations with IQ test scores on the Stanford-Binet at years 3 and 6 and the WISC administered as part of this investigation.

It was quite striking to find that there were fewer statistically significant correlations between the responses to cognitive demands at year 3 and the year 3 or year 6 IQ test results than with the WISC scores at the older age-period of the present study. Only three categories were significantly correlated with the 3 and/or

6 year IQ's: Verbal Not-Work responses - Competence positively and significantly correlated with the year and year 6 IQ data; Total responses expressed Verbally was positively related to year 6 IQ test scores; and Verbal Not-Work responses - Negation was negatively and significantly related to year 3 IQ test scores.

By contrast, six different categories of response to cognitive demands were significantly correlated with one or more scores on the WISC although this test was administered to most of the children about nine years after the year 3 "demands data" were observed. The following five categories were significantly and positively correlated with the Full Scale WISC IQ scores of the children: Total responses expressed as Work; Initial Work responses; Initial Not-Work responses followed by Work; Total responses expressed Verbally; and Work responses expressed Verbally.

These findings are in an expected direction since children who tend to work (rather than not-work) and to be verbal are likely to do better on the WISC.

The fact that these data involve a fairly long-term predictive analysis makes them quite impressive, especially since IQ stability over the same time span is relatively low. Moreover, the fact that the year 6 demands data yield only three significant correlations with the WISC (and these involve only two categories of response, neither of which correlated with the year 3 data) seems to suggest that the children's responsiveness to cognitive demands at an earlier age (three years) is more relevant to performance in middle childhood than is the cognitive response style at a later age (six years). In explanation, it may be that the year 6 data already reflect the children's adaptability to social expectancy rather than their more basic characteristics which are evident when they are younger. This may also account for the absence of significant correlations between the year 3 and year 6 demands data.

The study children's responses to the achievement motivation questionnaire showed them to have high

self concepts and sense of environmental control. This homogeneity was not surprising for a group of intellectually bright, advantaged children, but precluded the finding of a significant correlation with academic achievement.

Qualitative analysis showed that the behavior problem cases were distributed among the high, average and low achievers. The low achievers, with few exceptions, came from homes in which there was marked intrafamilial stress, either in terms of serious marital discord leading to separation or confused or disorganized functioning on the part of the mother. One exception was a stable family in which both children were low achievers. In this case, the parents were extremely permissive in their approach. The high achievers among the behavior problem cases were mostly from stable families, though here, too, there were a few exceptions in which marital discord played a role.

Outside of this issue of intrafamilial stress, the qualitative analysis of the behavior problem cases showed no other even suggestive correlations with academic achievement, such as with type and severity of problem or age at onset of problem.

RECOMMENDATIONS

The correlations found between the temperamental characteristics of the slow to warm up child (slow adaptability and withdrawal) and lower academic achievement confirm the significance of this constellation for school functioning indicated in one of our earlier reports (Gordon and Thomas, 1967). It would be important for teachers to be familiar with the characteristics of this child and the approaches desirable to maximize his learning and school achievement.

It is also important for school personnel to be aware that characteristics which are more obviously troublesome in the school setting, such as distractibility, high motor activity or those of the difficult child, do not necessarily predict unfavorable academic achievement.

The marked skewing of the scores on the achievement motivation questionnaire suggest the need for further research to define carefully the intragroup distribution of this variable in different socio-cultural samples.

References

1. Bridger, W. H. and Birns, B. M., 1963, Neonates Behavior and Autonomic Responses to Stress during Soothing, Recent Advances in Biological Psychiatry, V. New York: Plenum Press.

2. Brown, B., 1966, The Assessment of Self-Concept among Four-Year-Old Negro and White Children, presented at the Eastern Psychological Association Meeting.

3. Cattell, R. B., 1950, Personality: A Systematic and Factual Study. New York: McGraw-Hill.

4. Chess, S. and Korn, S., 1970, Temperament and Behavior Disorders in Mentally Retarded Children, Archives of General Psychiatry, 23:122-130.

5. Cohen, A., Linton, H. and Graham, ., 1959, Personality and Persuasibility, Vol. 2. New Haven: Yale University Press.

6. Cohen, J., 1965, Some Statistical Issues in Psychological Research. In Wolman, B. B. (Ed.), Handbook of Clinical Psychology. New York: McGraw-Hill.

7. Coleman, J. S. and Associates, 1966, Equality of Educational Opportunity. Washington, D. C.: Government Printing Office.

8. Crandall, V., Katkovsky, W. and Crandall, V. J., 1965, Children's Beliefs in their own Control of Reinforcements in Intellectual-Academic Achievement Situations, Child Development, 36:91.

9. Crandall, V., Katkovsky, W. and Preston, A., 1962, Motivational and Ability Determinants of Young Children's Intellectual Achievement Behaviors, Child Development, 33:643.

10. Escalona, S. and Heider, G., 1959, Prediction and Outcome. New York: Basic Books.

11. Finley, C. J., 1963, Comparison of the California Achievement Test, Metropolitan Achievement Test and Iowa Test of Basic Skills, California Journal of Educational Research, 14:79-88.

12. Gordon, E. M. and Thomas, A., 1967, Children's Behavioral Style and the Teachers' Appraisal of their Intelligence, Journal of School Psychology, 5:292.

13. Guilford, J. P., 1959, Personality. New York: McGraw-Hill.

14. Hertzig, M. E. and Birch, H. G., 1971, Longitudinal Course of Measured Intelligence in Preschool Children of Different Social and Ethnic Backgrounds, American Journal of Orthopsychiatry, 41:416-426.

15. Hertzig, M. E., Birch, H. G., Thomas, A. and Mendez, O. A., 1968, Class and Ethnic Differences in the Responsiveness of Preschool Children to Cognitive Demands, Monographs of the Society for Research in Child Development, Serial No. 117, Vol. 33, No. 1.

16. Kagan, J. and Moss, H., 1962, Birth to Maturity. New York: John Wiley and Sons.

17. Katz, I., 1967, The Socialization of Academic Motivation in Minority Group Children, Nebraska Symposium of Motivation.

18. Kay, S. A., 1967, The Relationship between Experiential Background and Individual Differences in Achievement among 4th and 7th Grade Pupils, Dissertation Abstracts, 27A:4133.

19. Lefcourt, H. M., 1966, Internal versus External Control of Reinforcement: A Review, Psychological Bulletin, 65:206-220.

20. McGhee, P. E. and Crandall, V. D., 1968, Beliefs in Internal-External Control of Reinforcements and Academic Performance, Child Development, 39:91-102.

21. Murphy, L. B. and Associates, 1962, The Widening World of Childhood. New York: Basic Books.

22. Pinneau, S. R., 1961, Changes in Intelligence Quotient. Boston: Houghton-Mifflin.

23. Raph, J., Goldberg, M. and Passow, 1966, Bright Underachievers. New York: Teachers College Press.

24. Tait, A. T., 1955, A Comparative Study of Five Major Achievement Tests, California Journal of Educational Research, 6:99-106.

25. Taylor, E. A. and Crandall, J. H., 1962, A Study of the Norm-Equivalent of Certain Tests Approved for the California State Testing Program, California Journal of Educational Research, 13:186-192.

26. Thomas, A., Chess, S. and Birch, H. G., 1968, Temperament and Behavior Disorders in Children. New York: New York University Press.

27. Thomas, A., Chess, S., Birch, H. G., Hertzig, M. E. and Korn, S., 1963, Behavioral Individuality in Early Childhood. New York: New York University Press.

28. Wechsler, D., 1951, Equivalent Test and Mental Ages for the WISC, Journal of Consulting Psychology, 15:381-384.

SUMMARY

The major aim of this study was to investigate the relationship between aspects of a child's behavioral style, specifically his temperament and task orientation, in the preschool age period and his level of academic achievement in elementary school. The relationship between behavior problems in the preschool or early school years and later academic achievement was also studied. The subjects in the study were 136 children of middle- and upper-middle-class background whom we have followed longitudinally from their first few months of life.

A great deal of data were available at the start of this project from our previous studies of these children. They included a behavioral description of each child in terms of his temperamental characteristics at ages 1 through 5; an analysis of each youngster's task orientation, based on his responses to the items on IQ tests administered at ages 3 and 6 years; IQ scores for each child at 3 and 6 years; and a clinical psychiatric evaluation of those youngsters who developed behavior problems.

In the present study, new data were collected on the children's academic achievement, intellectual functioning and achievement motivation. Each school a child had attended was contacted and requested to send us the youngsters' scores on all achievement tests he had taken. For purposes of analysis, the grade-point scores on tests taken in Grades 1 - 3 and 4 - 6 were pooled separately and then converted to percentile scores based on national norms. This was done for reading and arithmetic separately. In addition, the research staff psychologist individually administered the Wide Range Achievement Test (WRAT) during the testing session.

At this same testing session, each child was given the Wechsler Intelligence Scale for Children so that his current intellectual functioning could be determined. These scores, too, were converted to percentile scores based on national norms.

Finally, the youngsters were asked to complete an achievement motivation questionnaire devised to measure their self concept and sense of environmental control.

The quantitative analysis of the data emphasized correlational analyses between temperament and academic achievement, temperament and IQ, IQ and academic achievement, the temperamental characteristics of the "difficult" child and academic achievement (since these signs are associated with the development of behavior disorders), task orientation and IQ, task orientation and academic achievement and achievement motivation and academic achievement.

The temperamental characteristics of non-adaptability and withdrawal showed the most significant correlations with academic achievement. As a group, children with these characteristics do less well on measures of academic achievement to a significant degree.

Although temperament at year 5 was not related to IQ scores at ages 3 or 6, five of the nine characteristics were significantly related to the WISC Verbal Scale and, of these, four were significantly related to the WISC Full Scale score. These were approach, adaptability, high threshold and mild intensity.

The IQ scores at year 3 provided the earliest base for "predicting" later academic achievement and it was of interest that their correlations with achievement scores in Grades 4 - 6 were higher than those with scores in Grades 1 - 3 which were closer in time. Although the WISC scores cannot be viewed as predictive of achievement in Grades 1 - 3 or 4 - 6 since they follow the administration of those tests, these too were significantly correlated with the school measures of academic achievement and the WRAT.

Of special interest was the fact that children with signs of the difficult child, though at greatest risk for the development of behavior problems, did not appear at risk for academic underachievement. There was a lack of statistically significant correlation between this

temperamental constellation and the various measures of academic achievement. Similarly, the characteristics of high activity and distractibility, which have been known to interfere with a child's optimal functioning in school did not, in this group of youngsters, show any striking correlations with academic achievement.

The analysis of the relationship between task orientation and academic achievement suggested that the children's responsiveness to cognitive demands at 3 years was more relevant to performance in middle childhood than was their response style at 6 years. The greatest number of correlations were found between task orientation at 3 years and the WISC scores obtained on tests administered about nine years later.

The children's scores on the questionnaire assessing their self-concept and sense of environmental control were markedly skewed toward the high (positive) end. The average scores for the group ranged from 2.5 to 2.7 for self-concept and 2.7 to 2.9 for sense of environmental control; 3 was the high score and 1 the low in both cases. Probably as a result of this homogeneity, there was a lack of significant correlation between these measures of achievement motivation and academic achievement in this group of children.

Qualitative analysis of the behavior problem cases revealed a tendency for the low academic achievers in this group to come from homes in which there was severe intrafamilial stress. No other correlations were evident.