DOCUMENT RESUME

ED 079 414 TM 003 018

AUTHOR Ward, William C.
TITLE Disadvantaged Children and Their First School Experiences. ETS-Head Start Longitudinal Study: Development of Self-Regulatory Behaviors.
INSTITUTION Educational Testing Service, Princeton, N.J.
SPONS AGENCY Office of Child Development (DHEW), Washington, D. C. Bureau of Head Start and Child Service Programs.
REPORT NO ETS-PR-73-18
PUB DATE May 73
NOTE 72p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Age Differences; *Cognitive Development; Cognitive Processes; *Conceptual Tempo; Disadvantaged Youth; Longitudinal Studies; *Preschool Children; *Reaction Time; Self Control; Sex Differences; Technical Reports; *Test Validity
IDENTIFIERS Cognitive Style; Project Head Start

ABSTRACT

A three-year longitudinal study was conducted with 895 Head Start children to examine the development of self-regulatory abilities during the preschool years. The purpose was to discover, given the behaviors measured, whether there is convergent and discriminant validity for the existence of one or more dimensions of self regulatory behaviors during this period. Two cognitively-based measures of self regulatory behavior, the Matching Familiar Figures Test and the Motor Inhibition Test, were administered. Results are given in terms of levels of performance on the self-regulatory measures, internal characteristics of the scores from the MFF and MIT, interrelations among the self-regulatory behaviors, and their discriminant validity with respect to general ability and response tempo dimensions. Results indicate a change with age in the preschool years in the meaning of the self-regulatory behaviors. MFF latencies develop significant correlations with other scores as age increases. MFF errors showed a significant sex difference, with males making more errors than females each year. (KM)

PR-73-18

DISADVANTAGED CHILDREN AND THEIR FIRST SCHOOL EXPERIENCES

ETS-Head Start Longitudinal Study

Development of Self-Regulatory Behaviors

William C. Ward

May 1973

EDUCATIONAL TESTING SERVICE
PRINCETON, NEW JERSEY

PR-73-18

DISADVANTAGED CHILDREN AND THEIR FIRST SCHOOL EXPERIENCES

ETS-Head Start Longitudinal Study

Development of Self-Regulatory Behaviors

William C. Ward

Report under

Grant Number H-8256

Prepared for: Project Head Start
Office of Child Development
U. S. Department of Health,
Education and Welfare

May 1973

Table of Contents

Acknowledgments

The ETS-Head Start Longitudinal Study has entailed the efforts of well over a hundred individuals in the course of the last seven years. Many of these contributed much more of their time and energy than could reasonably have been asked, and are as deserving of thanks as those whose efforts are more recent and more visible.

Greatest appreciation is due to Virginia Shipman. Her leadership and unreasonable devotion, first as Coordinator for Measurement and Training, and for the last three years as Project Director, have been critical for the success, if not the survival, of the enterprise.

In the preparation of this report, Walter Emmerich, Robert Linn, and Virginia Shipman have provided thoughtful advice and helpful criticism. Appreciation is due also to Joan Tyson and her coding staff; to John Barone, Norma Hvasta, and Judith Ohls for their efforts in data analysis and programming; to Susan Simosko and James Towery for assistance in putting the pieces together; and to Thelma Benton and Gay Marshall for their typing skills and unflappable personalities.

Tables

Introduction

The development of control over one's own behavior has been a topic of central interest to research on cognitive abilities in the young child. White (1965) listed over 20 different phenomena showing more or less abrupt transitions in behavior during the age range 5 to 7, and offered an explanation linking these changes to a developing ability to withhold initially available responses long enough to permit a conceptually-based, more contingent response to be formulated. Flavell and his collaborators (e.g., Corsini, Pick, & Flavell, 1968; Moely, Olson, Halwes, & Flavell, 1969) have shown in examinations of children's learning and memory that in many cases age differences derive not from cognitive (e.g., mediational) deficiencies in the younger child, but from his lack of knowledge of appropriate strategies with which to approach a task. Young children can be made to perform more like older ones by forcing upon them the task approach spontaneously used by older ones; thus, the older child's superiority can be described in terms of greater ability to direct his task-relevant behavior according to an efficient strategy. From another perspective, Luria (1960) and more recently a number of American investigators (e.g., Wozniak, 1973) have been concerned with the development over age of the ability to use one's own verbalizations in regulating one's behavior. As a final example, the Piagetian progression of developmental periods may be conceptualized as movement from, initially, the development of c abilities in responding to the perceptually present world, toward both the ability to deal with an immediate situation without overt response, and the ability to deal with possible, rather than present, reality; i.e., toward behavior which is not constrained by immediate situational variables.

A similar emphasis may also be found within the realm of personality development. In psychoanalytic developmental theory, the transition from "primary process" to "secondary process" thinking is driven by the need to delay behavior leading to immediate rewards, in favor of more appropriate later rewards. More generally, much of development is conceptualized as the formation of defense mechanisms, or in neoanalytic theory of "cognitive control" mechanisms, which are structures providing the individual with means to be less immediately responsive to situational presses and more controlled by his own plans and desires.

A major question for the development of self-regulatory abilities concerns the degree to which these abilities here to form a dimension of individual differences. On the one hand, it is possible that alternative measures of self-regulation will correlate significantly with one another, allowing the description of an individual in terms of his level on a single dimension of self-regulation. On the other it may be that, while most individuals follow a course of increasing levels of control in general, different measures are only minimally related to one another at any given age--so that, rather than a single dimension of self-regulation, several dimensions are needed to describe the individual's standing at any one point in time.

The present report is concerned with the development of self-regulatory abilities during the preschool years. Its intent is to describe these behaviors in terms of levels, interrelations with one another, and relations with other measures of cognitive functioning. It is aimed at discovering, given the behaviors measured, whether there is convergent and discriminant validity for the existence of one or more dimensions of self-regulatory behaviors during this period. In several respects, it is a preliminary examination of this area. First, the age range examined extends from approximately 3 1/2 years of age to 6 1/2 years. It i- entirely possible that changes in both the dimensionality and the implications of these abilities continue to occur after this period. A later report will continue the examination of this domain into the early elementary-school years, and will permit further study of possible changes associated with developmental level and experimental diversity. Second, examination is limited to several measures that previously have served to represent the self-regulatory domain in the research literature, along with representative cognitive measures which may hel in interpreting nese behaviors. A later report dealing more complexly with self-regulatory behaviors and examining their relations to a broader set of cognitive and noncognitive measures may add to the understanding of the meaning and implications of these behaviors.

The unique contribution of the present study concerns not the specific self-regulatory measures employed, but rather the sample itself and the circumstances of administration. The sample is longitudinal, permitting conclusions about developmental changes which are unconfounded by sample differences across age. For the present age range, such differences can be more serious than they are in studies with older populations, when essentially all children of a given age will be in school, because of biases in which children are likely to be available for testing in any preschool year. In addition, the sample covers a broader range of socioeconomic level and preschool educational experience than is to be found in most studies. While it was drawn in such a way as to insure that a large proportion of the children would be economically disadvantaged, all children within a given school district were included in the testing once the district itself was selected. The result is a sample in which approximately 13% of the children had mothers with some education beyond high school and so might reasonably be classified as "middle class." Likewise, while selection was intended to assure that a large proportion of the children would attend Head Start programs, in the present sample 49% of the children attended Head Start, 21% attended other preschool programs, and 30% attended no program.

The sample is also large enough to permit detailed examination not only of main effects and mean levels, but also of correlations and interactions. For example, possible correlational differences associated jointly with sex and socioeconomic status are considered. By contrast, the large majority of studies in this area have worked with groups too small to allow more than at most a dichotomization according to a single demographic variable, usually sex.

These measures were collected as part of a large battery of cognitive tests, thus permitting eventual examination of the implications of self-regulatory behaviors over a broad range of differential ability measures. Beyond this there will also be the opportunity in subsequent reports to examine relations of self-regulation to more detailed information on the child's home environment, on his personal-social behaviors in the preschool, and on a variety of classroom behaviors.

The following section provides a description of the measures to be examined and the rationale for their inclusion. The tasks are characterized rather than being presented in detail. For a thorough description of these tasks, the reader is referred to technical reports covering the first two years of data collection in the larger study (Shipman, 1972c). Differences in the tasks as given in the third year of the study and comparisons of the present results with work by other investigators will be introduced in the text as appropriate.

Similarly, the subsequent sample description and outline of data collection procedures will be limited to summary information. Detailed presentations for the first two years of data collection can be found in Shipman (1972b, Chapters 2 and 3) and in briefer form in Meissner and Shipman (1973).

The final two sections of the present report involve a descriptive presentation of results and a discussion which will focus on the major points of importance in the present data and, in that context, on the relations of the present findings to results which have been presented by other investigators.

Measures and Rationale

Two cognitively-based measures of self-regulatory behavior provide the focus for this investigation. One of these, the Matching Familiar Figures Test (MFF), is a matching-to-sample task in which the child must indicate on each trial which of several forms, all but one differing in some small detail from a standard, is identical to that standard. Kagan has shown that on this task some individuals typically have long response latencies and low error rates, while others have short latencies and high error rates. The former are described as "reflective," the latter as "impulsive" (Kagan, Rosman, Day, Albert, & Phillips, 1964). These stylistic preferences are relevant to the self-regulatory domain in that the reflective individual is assumed to be using his longer delay period for more thorough processing of the information available to him; that is, he is choosing to withhold his response up to the point at which he will be able to respond with high accuracy. It is crucial to this interpretation that latencies and errors show the negative relation which Kagan and others have found. Latencies uncorrelated with quality of performance presumably would not represent such self-directed choice as to the necessary time for adequate information processing, but a simpler and less cognitively-relevant response tempo.

The second self-regulatory measure is the Motor Inhibition Test (MIT) (Maccoby, Dowley, Hagen, & Degerman, 1965). Here the child is given several simple motor acts to perform. After practicing each act, he is instructed to perform it as slowly as possible. The longer the time taken on "slow" administration trials, the greater the child's ability to modulate his response.

The use of these two measures to represent cognitive self-regulation is historically supportable--these are the two measures which have most often been used with young children, and so their use facilitates comparison of the present results with those of other investigators. In addition, they provide a possibly important contrast between stylistic and ability measures of self-regulation. On the Matching Familiar Figures Test the child is instructed to find the correct answer, without reference to speed of response; thus, his latencies represent his own decision as to how to proceed with the task, and allow him to demonstrate a preference for fast or slow performance. The Motor Inhibition Test, by contrast, has no performance criterion other than time. The child's attention is focused on the rate at which the response is given, providing a more direct measure of ability to regulate behavior.

Also included in the battery is a measure of the child's willingness to delay gratification (Mischel, 1958). The child is offered a choice between a small immediate reward and a larger delayed one; choice of the delayed reward is conceptualized as showing that the child has the ability to overcome the desire for immediate gratification. This performance apparently involves both personality and cognitive components, the former in that the child's capability to control expression of his desires is implicated, the latter in that he must be able to conceptualize a future reward as a justification for current self-denial.

In examining the dimensionality of self-regulatory behaviors, evidence on discriminant validity is as important as that on convergent validity. That is, one question is whether different regulatory behaviors are correlated with one another. A second is whether any relation they do have depends on variance specific to such behaviors, or whether these behaviors are related only as part of a larger dimension of cognitive development. Of chief concern is to discover whether a general ability or competence dimension includes such behaviors as one of its components. If this were the case, there would be little basis for discussing self-regulation as though it represented a distinct dimension of ability. For example, one might seek to discover family background characteristics associated with the rate at which the child progresses through a set of developmental stages, but one would not look for differential background characteristics specifically responsible for the development of self-regulation.

Three ability measures were examined in order to determine the degree to which self-regulatory behaviors are distinguishable from general ability level. One of these, the Preschool Inventory, is a standardized instrument sampling a broad range of skills and knowledge including general information, ability to follow directions, form reproduction ability, and understanding of several language and quantitative concepts. Though more explicitly dependent on specific learning opportunities than the general intelligence tests frequently used with older children, it would be expected to provide the same sort of general competence index. The other tasks had more focused content. The ETS Matched Pictures Test, modeled after the comprehension task employed by Fraser, Bellugi, and Brown (1963), provides an assessment of the child's knowledge of grammar; he must show his understanding of a given grammatical contrast by choosing from two similar pictures the one that is appropriately labeled with on ple of the contrast (e.g., "One of these pictures is called <u>Bear is sitting</u>, and one is called <u>Bear is not sitting</u>. Show me the picture called <u>Bear is not sitting</u>.")

The third test is the ETS Enumeration Test. The child is shown a page of colored circles, with pages varying in the number, coloring, and systematicity of the array, and is required to point to each circle once and only once. The test requires the child to establish a one-to-one correspondence between his pointing and the circles on the page, an ability which in Piagetian theory is one of the precursors to the development of quantitative concepts.

Since the two major self-regulatory tasks involve the measurement of ed of response, it is also important to discover whether any variance they share is attributable to a more general dimension related to response tempo. Their conceptualization in self-regulatory terms involves the assumption, in both tasks, that long latencies are produced in the service of a cognitive goal, not because the subject is unable to respond quickly. This distinction is particularly important in understanding "impulsive" performance. Some individuals, for example, may show "impulsivity"--i.e., quick response, high distractibility, etc.--for physiological reasons rather than cognitive ones; the implication of such fast responding would be different from impulsivity as conceptualized here. In order to have some basis for determining whether a general response tempo dimension contributes to any relations among time measures on the self-regulatory tasks, scores from the several additional tasks which provided measures of response latency are also considered. These were latencies to first object choice on the Sigel Object Categorization Test, latency to first response on the Preschool Embedded Figures Test, and fastest time to correct completion of object placement on the Seguin Form Board.

Failure of time scores from these tasks to relate to the self-regulatory measures would provide evidence that variance on the latter is specific to the self-regulatory domain. On the other hand, a finding of uniformly positive relations might be ambiguous. Each of these three task situations contains some of the elements which Kagan has argued provide the necessary context for elicitation of variance attributable to the reflective-impulsive style. The Sigel Object Categorization Test provides the strongest case, in that, as required by this argument, there are a number of simultaneously available alternatives (the 12 objects) from which the child must make a selection, along with uncertainty as to what should be chosen. However, performance of young children on this task suggests that the number of choices they consider is much smaller; those who succeed in making a defensible grouping most often do so by use of a single obvious perceptual characteristic of the object, usually color. For such groupings, the child may not in fact be operating with any response uncertainty; a better correspondence to Kagan's required conditions would require the child to see a choice between, for example, grouping by form or by color. On the Preschool Embedded Figures Test, the child is instructed to find a standard triangle embedded within a complex visual scene. Again, it is possible that he is faced with a choice among several alternative possible locations of the triangle, and that his response delay is related to his stylistic preference for a greater or lesser certainty. Here, though, even more than in the sorting task, it seems unlikely that alternatives occur simultaneously; they are more likely to become apparent in succession, and to depend on the child's level of competence in disembedding more than on his stylistic preferences. Finally, the Seguin Form Board task requires the child to place 12 blocks of different shapes within their corresponding recessed niches in a board; the time measure is total time taken to successful completion of the task on the fastest of the three trials the child is given. Here, perceptual-motor coordination rather than self-regulatory behaviors appear to be a more relevant explanation of the performance. Two of these tasks, therefore, seem to involve time scores which are not easily interpreted as relevant to the self-regulatory dimension. Only the Sigel Object Categorization Test can very plausibly be interpreted in these terms, and for this task the rationale is more convincing at older ages, when it is reasonable to presume that children are aware of and must decide among several alternative grouping rules, than during the early preschool years.

Method

Sample Characteristics

Four regionally distinct communities were selected which 1) had a sufficient number of children in school and in the Head Start program, 2) appeared feasible for longitudinal study given expressed community and school cooperation and expected mobility rates and 3) offered variation in preschool and primary grade experiences. The study sites chosen were Lee County, Alabama; Portland, Oregon; St. Louis, Missouri; and Trenton, New Jersey. Within these communities, elementary school districts with a substantial proportion of the population eligible for Head Start were selected for participation. In each school district an attempt was made to test all children of approximately 3 1/2 to 4 1/2 years of age in the initial testing and data collection of 1969, although some children were excluded from the sample (e.g., children from families whose primary language was not English, and those with severe physical handicaps).

In 1969 mothers were interviewed and children tested prior to their enrollment in Head Start or any other preschool program. The following is an overview of the salient demographic characteristics of the initial four-site sample (for a more complete description of this population the reader is referred to Project Report 71-19 [Shipman, 1971]):

1. At least partial data were obtained for a total of 1875 children. However, the number of subjects at each site varied, with Lee County and Portland together constituting 60% of the sample.
2. The sample was 62% black.
3. Boys made up 53% of the sample. For the four sites they comprised 54.5% of the black sample and 50.5% of the white sample.

4. For the three sites in which children had the opportunity to attend Head Start in Year 2 of the study (1969-70), 37.2% of the sample attended Head Start, 11% attended other preschool programs and 51.8% had no known attendance in Head Start or other preschool programs. In Lee County, where Head Start is a kindergarten level program, 41.7% of the initial sample attended Head Start, 19.1% attended other preschool programs and 39.3% had no known attendance in Head Start or other preschool programs.

5. Substantially more blacks than whites attended Head Start. While this varied by site, in the total sample, only 5.1% of the children who attended Head Start were white.

The Year 2 sample included children from four sites: Trenton, Lee County, Portland, and St. Louis. During Year 2 data-gathering procedures in Lee County were limited to a portion of the test battery due primarily to limited resources and to the fact that most of the children in Lee County were not enrolled in preschool programs until the third year of the study.

The Year 3 sample included children from three sites; no individual testing took place in St. Louis, due to a combination of problems in field operations and loss from the first to the second year of a substantial portion of the sample, which would have resulted in a longitudinal sample too small to justify the investment of resources required for testing. Since this was the Head Start year in Lee County, the most extensive data collection was done with children in this site; those in Portland and Trenton, for whom this was the kindergaten year, were given abridged test batteries.

For the present report, the sample employed in all analyses was limited to those children who were given at least part of the testing battery during all three years. In general, the measures included in this report were given in the reduced batteries in Lee County in Year 2, and in Portland and Trenton in Year 3. Thus, analyses include children from three testing sites; in most cases the N for a given measure approaches 895, the total three-year three-site sample. Actual N's are somewhat smaller than this figure, since a measure may not have been given to a particular child in a particular year, or if given may have been invalidated due to child failure to meet a criterion on practice items or to tester error in administering or recording. This type of variability in sample was considered more acceptable than the several alternatives. For example, limiting the sample to children whose data were complete in all respects would have reduced the available number drastically and in a way which may have been contingent on the child's ability level.

A detailed breakdown of sample characteristics is given in Tables 1 to 3. Table 1 combines children from the three testing sites and presents a four-fold cross-classification of the sample. Children are grouped according to mother's education which serves as an index of socioeconomic status (less than 10 years, 10-12 years, more than 12 years), child's sex, race, and preschool attendance. For the latter classification, the categories included attended Head Start, attended another preschool program, and (so far as study records show) attended no preschool program. This classification depends on the child's preschool attendance in the Head Start year in each locality. In Trenton and Portland, the Head Start year was Year 2, and in Lee County it was Year 3.

Table 1

Total Sample Classified by Sex, Race, SES, and Preschool Attendance

Males-Low SES*

	HS	PS	NONE	
Black	72	6	28	106
White	3	1	7	11
	75	7	35	117

Females-Low SES

	HS	PS	NONE	
Black	60	5	16	81
White	4	0	9	13
	64	5	25	94

Males-Middle SES*

	HS	PS	NONE	
Black	127	30	61	218
White	11	30	31	72
	138	60	92	290

Females-Middle SES

	HS	PS	NONE	
Black	120	18	45	183
White	12	20	37	69
	132	38	82	252

Males-High SES*

	HS	PS	NONE	
Black	7	3	1	11
White	2	28	8	38
	9	31	9	49

Females-High SES

	HS	PS	NONE	
Black	7	6	2	15
White	2	36	12	50
	9	42	14	65

* SES is defined by mother's education:

Low SES - Less than 10 years education
Middle SES - 10-12 Years education
High SES - More than 12 years education

** HS - Attended Head Start
PS - Attended other preschool program
NONE - No known preschool attendance

In general, the three-year three-site longitudinal sample was demographically similar to the initial one-year sample. This sample was 53% male and 71% black; as in the initial year, a slightly higher proportion of black children (55%) than white children (48%) were male. Forty-nine percent of the longitudinal sample attended Head Start, 21% attended other preschool programs, and 30% attended no preschool program. Comparison with the Year 1 percents suggests that the project was somewhat more successful in maintaining contact with those children who participated in some preschool program than with those who did not. As in the initial sample, the Head Start population was largely black (92% as opposed to the earlier 95%), while the other two attendance categories showed a more even racial division.

Table 2 presents single breakdowns of the data separately by testing site. The only notable differences between sites are 1) while the sample was largely black in all sites, the percent of whites was much smaller in Trenton than in the other two sites; 2) in Lee County, a larger percent of the children than in the other two sites attended preschool programs other than Head Start, while a smaller percent attended no preschool programs; and 3) the mean socioeconomic status, as indexed by mother's educational level, was higher in Portland than in the other two sites. More detailed breakdowns of the sample characteristics for each site (not presented in the tables) show that there was no sex by race or by preschool attendance disproportionality in any of the sites, and that those who attended Head Start were largely black in all sites. There was a race by preschool attendance disproportionality in that both Lee County and Portland included a number of white children who attended preschool programs other than

Table 2

Demographic Characteristics of Sample
Classified by Site

Total N	Trenton 226	Lee County 391	Portland 278	3-Site Total 895
Males	119(52.6%)	213(54.5%)	142(51.1%)	474(53.0%)
Females	107(47.4%)	178(45.5%)	136(48.9%)	421(47.0%)
Black	203(89.8%)	240(61.4%)	192(69.1%)	635(71.0%)
White	23(10.2%)	151(38.6%)	86(30.9%)	260(29.0%)
Head Start	101(44.7%)	201(51.4%)	142(51.1%)	444(49.6%)
Other Preschool	28(12.4%)	127(32.5%)	31(11.1%)	186(20.8%)
None	97(42.9%)	63(16.1%)	105(37.8%)	265(29.6%)
Mother's Education:				
Mean	10.32	10.70	11.55	10.87
SD	0.41	0.49	0.38	0.46

Head Start (98 of 151 whites in Lee County and 17 of 86 whites in Portland), while this cell was nearly empty in Trenton (1 of 23 whites).

The various demographic imbalances and disproportionalities in sample characteristics are a necessary consequence of the subject selection procedures initially employed in the study and of our nonintervention in the preschool educational decisions made for each child. Moreover, any attempt to create a more balanced sample would have rendered the sample less representative of the preschool attendance groups actually existing in study communities. For interpretive purposes, the major difficulty raised by these imbalances concerns the confounding in the sample of race, socioeconomic status, and preschool experience. Head Start in this sample was largely black, while other preschool programs were attended by children of both races; black children were drawn more often from lower SES homes, and whites from the upper end of the SES range. The disproportionalities were too great to permit simultaneous study of race in conjunction with either of these variables: the cells representing whites of the lowest SES status, or blacks of high SES, for example, are essentially empty. An alternative to combining races would have been a limitation of the sample to blacks only, but this would have restricted the SES range severely. In the present report, as has been the practice in earlier reports from the project, the decision has been to focus attention on subject differences which appear most likely to yield information relevant to process differences; thus, analyses are done considering performance as it relates to and is affected by socioeconomic status, rather than by race of child. An earlier series of analyses of Year 1 data (Shipman, 1971) provides additional support for this approach, in that no race difference in the factor structure of the larger cognitive battery was found; in addition, analyses of parent interview data (Shipman, 1972a) suggest that the meaning of the SES indicator employed here, mother's educational level, is similar across race.

Data Collection

Field offices were established in the participating communities and, so far as possible, local personnel were given responsibility for administrative and data collection activities in their own community. During the first year of data collection, ETS staff took major responsibility for these activities, including training and evaluating testers, frequent monitoring of testing, and solving problems that arose in the course of these activities. In subsequent years of operation, local coordinators assumed a larger share of administrative responsibilities, while technical advisors took more responsibility for tester training and for monitoring.

Local women were trained as testers. The usual educational credentials were not required; selection depended on the ability to learn to administer a battery of tests and on the quality of interaction with young children. Most of the testers were black housewives with limited work experience. Testers were given a training period of four to six weeks, followed by a "dry run" operation of the testing centers for an additional one or two weeks before actual data collection began.

Instruments were arranged in batteries, balanced to provide variety in task format and domain measured, and approximately equal levels of task difficulty for both testers and children.

Testing centers were located in churches or community recreation facilities in or near the districts where the children lived[1]. Each center provided at least six individual testing rooms or partitioned spaces and a large play and rest area; most also included kitchen facilities. Each center, operating five days a week, was staffed by nine persons--a center supervisor, a play-area supervisor, a driver, and six testers.

During the first year of data collection, children were scheduled for a four-day testing sequence with each session lasting about 1 1/2 hours. The first day's testing involved three mother-child interaction tasks plus several individual measures; for the remaining three days, the

---

[1] In Year 3, some data collection took place in the children's schools rather than study test centers.

batteries of individual tests remained intact but different children were given them in all possible inter-day sequences. The child was tested by the same tester on the second as on the first-day's session, and subsequently by two different testers to complete the series. In the second and third years of data collection, a three-session series was given, with the child seeing a different tester on each day, and again with different children receiving the three batteries in different sequences. As indicated earlier, in Lee County in Year 2, and in Trenton and Portland in Year 3, only a subset of the individual tests were given; in these cases the child received a 1 1/2-hour battery in one day.

Testing schedules remained quite flexible. Children were tested only when they were deemed able to cooperate and to show their ability; when necessary they were given rest periods in the play area or were brought back to the testing centers on additional days.

A listing of the composition of the complete test battery for each of Years 1 and 2 may be found in earlier project reports (Shipman, 1972b). Most tests remained identical over these two years, though refinements in manuals, amounting mainly to further clarifications for the tester, were made. A number of changes were made for the third year of testing--increasing the difficulty of the test by the addition or substitution of items, dropping measures which had become too easy or which did not appear to be yielding useful information, and adding measures which previously had been too difficult at the younger ages.

The measures here under study remained constant over the three years, except for some item changes intended to increase difficulty level. In Year 3, for example, the last 10 items of the MFF were replaced by 10 more difficult items. In the same year, four additional trials were appended to the Motor Inhibition Test to provide an assessment of the child's ability to perform each of the motor acts as fast as possible. The Enumeration and Matched Pictures Tests were changed more drastically with the addition of new item types; thus, changes in their correlates may be in part a function of test-content changes. The other measures considered here remained constant over the three-year period.

The tasks considered in this report are listed in Table 3, along with the number of children for whom scores were obtained on each task in each of the three years of data collection. The Seguin Form Board and the Mischel Delay of Gratification were not given in Year 3, nor was the latter given in Lee County in Year 2; the Preschool Inventory was given only in Lee County in Year 3; Matched Pictures and Enumeration were given only in Trenton and Portland in Year 2. Otherwise, each measure was given to all subjects in each year.

Table 3

N's for Measures for Each Year

Task	Year 1	Year 2	Year 3
MFF	828	871	892
Motor Inhibition	860	872	890
Mischel Technique	853	493	---
Sigel	661	453	861
PEFT	757	452	887
Preschool Inventory	855	871	389
ETS Matched Pictures	842	874	894
ETS Enumeration	814	792	858
Seguin Form Board	657	484	---

Results

In this section attention will be given to levels of performance on the self-regulatory measures, to internal characteristics of the scores from the MFF and MIT, to interrelations among the self-regulatory behaviors, and to their discriminant validity with respect to general ability and response tempo dimensions.

Levels of Self-regulatory Behaviors

Table 4 presents summary information on the self-regulatory measures over three years. For MFF the scores employed were 1) mean latency to first response on valid items of the first eight test items (those items common to the task as given in every year), and 2) mean errors for these items. To reduce effects on mean latencies of occasional very long latencies, which in young children would probably represent lack of attention rather than reflectiveness, the latencies were transformed by log (X+1) before averaging. The child was given two response options on each item permitting a maximum possible error score of 16 for these items.

Six scores were computed for MIT, four representing each individual trial--practice on the walking and drawing subtest and performing each of these as slowly as possible--plus two summary scores. Each individual trial score was transformed by log (X+1). The two summary scores were computed by standardizing the individual trial scores to a mean of 50 and standard deviation of 10, based on the mean and standard deviation obtained for the total four-site sample in Year 1. These standardized scores were then averaged over the two practice subtests to give a summary of the child's performance when instructed to perform slowly. It should be noted that the mean and variance which served as the basis for standardization were the

Table 4

Sample Means and Standard Deviations

on Self-Regulatory Measures

	Year 1			Year 2			Year 3		
Measure	Mean	SD	Mean in Seconds	Mean	SD	Mean in Seconds	Mean	SD	Mean in Seconds
MFF-Response Time	.594	.133	2.93	.617	.128	3.14	.640	.143	3.37
Errors	.682	.349		.533	.323		.324	.254	
MIT-Walk Practice	.755	.159	4.96	.744	.145	4.50	.723	.152	4.28
Slow	.868	.215	6.38	.938	.219	7.67	.999	.204	8.97
Draw Practice	.669	.210	3.67	.637	.186	3.34	.747	.209	4.58
Slow	.835	.291	5.84	.949	.293	7.89	1.125	.282	12.34
Average Practice	49.661	7.665		48.208	7.314		50.337	7.472	
Average Slow	49.837	8.936		53.375	9.141		57.992	8.302	
Mischel	.660	.474		.663	.482				

parameters obtained for the total four-site sample in Year 1 testing. Use of these parameters served to permit summary scores in which the two subtests contributed equally to the summary, but which preserved information as to mean differences over years.[2]

Looking first at the MFF, two observations are obvious. First, latencies increased over years, while errors decreased. Analysis of variance shows the change over years to be a highly significant ($p < .0001$) linear trend for each measure. These changes are in accord with results reported by Kagan (1966) and Kagan et al. (1964). The lower error scores with increasing age are predictable in that any ability measure is likely to show such a trend. Longer latencies with increasing age, however, suggest the operation of the stylistic variable, namely that older children take longer to process the information available to them, thereby making a more adequate assessment. Their superior performance may then be due in part to this increase in the thoroughness of their search of the information available.

Secondly, at all ages the latencies were very short; the most "reflective" children in this sample were responding quickly in absolute terms. Reconverting the latency means to seconds yields, for Years 1 through 3, respectively, means of 2.93, 3.14, and 3.37 seconds. An examination of the distribution of mean latencies reinforces the point. In the Year 2 four-site data, on which such an analysis was made, the central tendency was very similar to that for Year 2 data in the present sample: a mean equivalent to 3.01 seconds. Here 90% of the cases fell between means

---

[2]For MFF latencies and for the four individual trials on MIT, Table 4 also contains the equivalent scores in seconds along with the means and SD's of the log transformed scores.

of 1.97 and 5.96 seconds, the fifth and ninety-fifth percentiles of the distribution.

Turning to the Motor Inhibition Test trials, it is evident that at all ages the instruction to perform an act slowly was understood and followed by the subjects. The differences between practice and slow performance on each subtest at each age amounted to at least 0.7 standard deviations, expressed in terms of performance on practice trials. Expressed otherwise, when the log mean scores were reconverted to seconds, time taken to execute the act slowly ranged from 29 to 110% more than for practice trials on the walking subtest and from 59 to 169% more on the drawing subtest. More detailed analysis of these data was undertaken with the four-site Year 1 sample broken into six groups according to the age of the child at the time of Year 1 testing. Here, for both subtests, every age subgroup showed a difference in mean time; thus, even the youngest subjects tested in the study were able to respond to the request to slow their response execution.

Times for all four trials showed significant changes with age. For the walking subtest, the changes were linear--practice times decreased, while slow administration times increased ($\underline{p} < .0001$). For the drawing subtest, both scores showed significant linear ($\underline{p} < .0001$) and quadratic ($\underline{p} < .0002$) changes with years. For practice on the drawing subtest, time taken decreased from Year 1 to Year 2 and then increased in Year 3. It is possible that the decrease from first to second years represents an improvement in skill in the fine motor coordination required to draw a line using pencil and ruler. The trend for slow administration time for the drawing subtest, finally, showed an increase from each year to the next with the increase larger from Year 2 to 3 than from Year 1 to 2.

The difference in trend over time for both subtests, for slow administration versus practice trials, serves to confirm that the MIT is measuring self-regulatory rather than response tempo differences--older children slow their responses more than younger ones when asked to do so, but not (with any consistency) when instructions do not call for a slow performance.

The two summary scores for the task provide an adequate representation of the data. For the average practice score, there was no linear age trend, but a significant ($p < .0002$) quadratic trend. That is, mean performance showed a slight decrease in time from Year 1 to 2 and a slight increase from Year 2 to 3, the largest difference amounting to approximately 0.2 standard deviation. Average slow administration performance, however, showed a significant ($p < .0001$) linear trend and a marginal ($p < .04$) quadratic trend. Times increased about 0.4 standard deviations from Year 1 to Year 2 and over 0.5 standard deviations from Years 2 to 3.

Finally, the Delay of Gratification task yielded 66.0% of the subjects choosing the delayed reward in Year 1 and 63.3% choosing it in Year 2; the task was not given in Year 3. The age trend was not significant.

In order to examine the consistency of trends in self-regulation, a series of repeated measures analysis of variance was performed (sex x SES x year, with SES represented by three levels of mother's education--less than 10 years, 10-12 years, and more than 12 years). Means and standard deviations for each sex, each SES level, and each combination of these variables are presented in Table 5 for MFF latencies, in Table 6 for MFF errors, and Tables 7 and 8 for MIT average practice and average slow administration scores, respectively. No consistent pattern of differences

Table 5

MFF Latencies Log (X+1), Items 3-10
Means for Sex X SES Subgroups by Year

	Low SES Mean	Low SES SD	Middle SES Mean	Middle SES SD	High SES Mean	High SES SD	Total Mean	Total SD
MALE								
Year 1	.589	.135	.611	.141	.615	.121	.605	.137
Year 2	.607	.156	.606	.120	.627	.103	.609	.127
Year 3	.630	.122	.643	.154	.648	.154	.637	.146
FEMALE								
Year 1	.550	.126	.591	.126	.589	.124	.583	.127
Year 2	.643	.139	.617	.128	.636	.107	.626	.129
Year 3	.618	.154	.641	.130	.667	.143	.642	.140
TOTAL								
Year 1	.571	.132	.602	.135	.600	.123	.594	.133
Year 2	.623	.149	.611	.124	.632	.105	.617	.218
Year 3	.624	.137	.642	.143	.659	.147	.640	.143

Table 6

MFF Errors, Items 3-10

Means for Sex X SES Subgroups by Year

	Low SES		Middle SES		High SES		Total	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
MALE								
Year 1	.750	.375	.705	.334	.516	.343	.695	.348
Year 2	.691	.358	.544	.314	.344	.284	.558	.336
Year 3	.413	.294	.366	.271	.203	.192	.360	.274
FEMALE								
Year 1	.772	.364	.679	.341	.498	.305	.668	.349
Year 2	.621	.294	.500	.307	.354	.241	.505	.306
Year 3	.365	.251	.284	.216	.176	.168	.283	.224
TOTAL								
Year 1	.760	.369	.693	.337	.505	.320	.682	.349
Year 2	.661	.333	.524	.311	.350	.260	.533	.323
Year 3	.392	.276	.328	.250	.187	.178	.324	.254

Table 7

MIT Average Practice Score

Means for Sex X SES Subgroups by Year

	Low SES Mean	Low SES SD	Middle SES Mean	Middle SES SD	High SES Mean	High SES SD	Total Mean	Total SD
MALE								
Year 1	48.674	8.199	49.494	7.706	50.581	7.218	49.470	7.770
Year 2	47.070	7.331	47.886	7.037	49.120	5.851	47.822	7.032
Year 3	50.306	6.740	49.879	7.287	50.112	6.262	50.052	7.004
FEMALE								
Year 1	49.800	8.486	49.860	6.933	50.780	8.538	49.874	7.559
Year 2	47.992	7.815	48.661	7.315	48.789	7.976	48.643	7.606
Year 3	48.985	6.717	51.294	8.006	49.713	7.808	50.658	7.964
TOTAL								
Year 1	49.184	8.328	49.665	7.350	50.693	7.953	49.661	7.669
Year 2	47.469	7.539	48.247	7.171	48.929	7.123	48.208	7.314
Year 3	49.721	6.746	50.537	7.655	49.885	7.157	50.337	7.472

Table 8

MIT Average Slow Score

Means for Sex X SES Subgroups by Year

	Low SES		Middle SES		High SES		Total	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
				MALE				
Year 1	46.970	9.147	48.961	8.546	55.697	8.250	49.322	9.010
Year 2	51.371	8.820	52.225	8.654	58.766	9.315	52.789	8.946
Year 3	57.030	7.268	57.082	8.230	61.721	9.623	57.591	8.233
				FEMALE				
Year 1	47.852	8.140	50.692	8.947	53.961	8.272	50.411	8.839
Year 2	50.482	8.316	54.148	9.527	57.607	8.182	54.037	9.323
Year 3	56.316	7.848	58.526	8.060	60.902	9.616	58.446	8.366
				TOTAL				
Year 1	47.369	8.696	49.770	8.770	54.720	8.271	49.837	8.941
Year 2	50.984	8.594	53.122	9.114	58.099	8.660	53.375	9.141
Year 3	56.714	7.520	57.752	8.176	61.254	9.585	57.992	8.302

emerged from these analyses, though there were scattered significant effects in three of them.

For MFF latencies, there was a marginal sex by year interaction, with significant linear ($p < .04$) and quadratic ($p < .02$) components. Examination of Table 5 shows that males had slightly longer latencies than females in Year 1, while females had longer latencies than males in Year 2 and the groups were nearly equal in Year 3. The largest of these differences, however, amounts to only .12 standard deviations. There was also a year by SES interaction in the latency data (quadratic component significant, $p < .005$). In general, latencies increased with year, though the increase was small in absolute or percentage terms. For the high SES group, the increase was essentially linear, for the middle group it was more rapid from the second to the third year, and for the low group it was more rapid from the first to the second year.

MFF errors showed a significant sex effect ($p < .01$) with males making more errors than females in each year. In addition, there was a marginal age by year interaction (linear component significant at .04): the difference in error scores for males and females increased with age. Males averaged .03 more errors than females in Year 1, .05 more in Year 2, and .08 more, per valid item, in Year 3. In percentages, males made, respectively, 4, 10, and 27% more errors than females in Years 1, 2, and 3. MFF errors also showed a significant SES effect, with large error differences associated with SES level, and a year by SES interaction (quadratic component significant, at $p < .02$), which apparently resulted from a greater decrease in errors over years for children of the highest SES level than for those from the other two levels.

On the MIT, no sex or SES differences were found associated with average practice time. Average slow time showed a significant effect of SES ($p < .0001$) and a marginal ($p < .05$) interaction of sex and SES: low SES subjects showed no sex difference, middle SES subjects showed slight sex differences favoring females, and high SES subjects showed slight sex differences in favor of longer times for males. Since these differences average less than 0.2 standard deviations, they are not large enough to deserve serious effort at interpretation.

Results from the Delay of Gratification task were also examined; means and standard deviations are provided in Table 9. No systematic differences were found.

Internal Characteristics of MFF

In this section attention is given to the reliabilities and stabilities of the MFF scores and to the interrelation of the MFF latency and error scores. In contrast to the presentation given in the previous section, the scores used here and in subsequent analyses are mean transformed latencies and mean errors per item over all 18 test items administered in a given year; these are the most reliable indices available.

Coefficient alpha reliabilities are given in Table 10. Latencies had a consistently high level of internal consistency in all years. Errors had satisfactory but more modest reliability; the lower coefficient in Year 3 may indicate that some of the items were becoming too easy. These coefficients were obtained using data from the entire sample tested each year. Very similar values were obtained when the sample was broken in several ways, e.g., by site and, in more detailed examination of Year 1 data, by sex, race, age, or SES level. For the Year 1 subgroup analyses, for example, the reliabilities obtained ranged from .87 to .92 for latencies and from .63 to .76 for errors.

Table 9

Mischel Choice Score

Means for Sex X SES Subgroups by Year

	Low SES		Middle SES		High SES		Total	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
MALE								
Year 1	.664	.475	.656	.476	.739	.444	.671	.470
Year 2	.674	.474	.608	.490	.571	.507	.631	.483
FEMALE								
Year 1	.674	.471	.612	.488	.688	.467	.648	.478
Year 2	.581	.499	.627	.485	.704	.465	.634	.483
TOTAL								
Year 1	.668	.472	.636	.482	.709	.454	.660	.474
Year 2	.629	.486	.618	.487	.646	.483	.633	.482

Table 10

Reliabilities of MFF Scores

	Latency	Errors	N
Year 1	.90	.70	1404
Year 2	.91	.71	1304
Year 3	.92	.53	1157

Correlations over time among latencies and among errors are given in Table 11 for the total longitudinal sample, and by sex, SES level, and the various sex by SES level combinations. For the total sample, errors had reasonably high stability over time in relation to their reliabilities. The correlations corrected for attenuation were .68, .62, and .77 for Years 1 x 2, 1 x 3, and 2 x 3, respectively. Latencies, however, had minimal consistency over time despite their greater within-year reliability; correction for attenuation yields coefficients of .18, .20, and .21. Examination of the coefficients according to sex and SES breakdowns does not suggest any systematic difference in stability associated with these factors. Isolated comparisons can be made which would yield significant differences, but in the absence of specific hypotheses regarding which comparisons might show differences, these are likely to represent non-meaningful fluctuations rather than generalizable differences.

Correlations between latencies and errors within each year's data are presented in Table 12. For the total longitudinal sample, the coefficient is positive in Year 1 and nonsignificant (r = -.06, p < .10). It is low and negative in Year 2 (r = -.10, p < .005); and it is stronger

Table 11

Stability of MFF Scores Over Years for Sex X SES Subgroups

	Low SES Years			Middle SES Years			High SES Years			Total Years		
	1x2	1x3	2x3	1x2	1x3	2x3	1x2	1x3	2x3	1x2	1x3	2x3
Latencies												
Male	.09	-.04	.21	.27	.16	.22	.22	.36	.24	.21	.14	.21
Female	.15	.28	.07	.16	.18	.19	.05	.25	.48	.13	.24	.17
Total	.09	.13	.13	.21	.17	.21	.12	.28	.37	.16	.18	.19
Errors												
Male	.40	.35	.31	.42	.33	.38	.54	.53	.64	.47	.40	.42
Female	.40	.29	.45	.44	.32	.51	.46	.11	.27	.49	.34	.51
Total	.39	.32	.36	.43	.33	.45	.50	.33	.45	.48	.38	.47

and negative in Year 3 ($r = -.36$, $p < .0001$). The difference in correlations from year to year is highly significant; by a conservative test, the Year 1 and 2 correlations differ at the .001 level, while Year 2 and 3 correlations differ at the .0001 level. These correlations support the interpretation that in Year 1 response latencies did not serve as an index of a reflective or impulsive style as they did in Year 3. In the intermediate year some evidence for the presence of the style is found.

Table 12

Correlations of MFF Latencies With Errors
for Sex X SES Subgroups by Year

	Low SES			Middle SES			High SES			Total		
	Year			Year			Year			Year		
	1	2	3	1	2	3	1	2	3	1	2	3
Male	.06	.14	-.42	.12	-.06	-.28	.08	-.31	-.21	.09	-.01	-.30
Female	.10	-.09	-.36	-.02	-.29	-.43	.24	-.29	-.41	.01	-.20	-.43
Total	.07	.03	-.38	.07	-.17	-.34	.17	-.30	-.31	.06	[illegible]	-.36

The same relative sequence of correlational changes was found for each sex, each SES level, and for five of the six sex by SES subgroups. There are, however, two further differences of interest presented in Table 12. First, there was a sex difference in the correlations. In Year 3 females showed a marginally stronger negative correlation between latencies and errors than did males ($p < .05$). In Year 2 they had a significant negative correlation ($r = -.20$, $p < .0001$), where males showed a near zero relation ($r = -.01$); these coefficients differed significantly ($p < .005$). In Year 1 there was no significant sex difference in the correlations, though the coefficient for females was a lower positive correlation than that for males ($p < .25$). It appears, therefore, that the reflection-impulsivity dimension of self-regulation became a contributor to the behavior of the females earlier than it manifested itself in males, but that by the kindergarten year it was evident in children of both sexes.

Second, a comparable pattern was found in socioeconomic level. In Year 1 latencies and errors were uncorrelated for children of all SES levels and in Year 3 they were negatively correlated for all levels. In Year 2 they were uncorrelated for low SES children, had a low negative correlation for middle SES children, and were more strongly negatively correlated for the high SES group. Again, there appears to be a difference in rate of development of the stylistic variable, with its manifestation apparent earlier in children of higher socioeconomic status backgrounds. Finally, considering sex and SES together, the coefficients were as would be expected: for neither sex was there a negative relation in Year 1; for both there was such a relation in Year 3; and in Year 2 the relation was found for females of middle and high SES, but for males only in those of high SES.

Internal Characteristics of MIT

Table 13 presents reliabilities for the MIT average practice and average slow scores. These are coefficient alphas derived from the correlation of the two practice or the two slow administration subtest scores. The table includes, for Year 3, the comparable coefficients for practice and fast administration trials given after completion of the four trials which were common to all years' data. Practice scores were not strongly related to one another; the uncorrected correlations ranged from .16 to .22 for the initial practice trials. "Slow" trials (as well as "fast" trials) had a higher level of consistency.

Correlations of the average practice and of the average slow trials over years are given in Table 14 for the total longitudinal sample and for the various sex and SES subgroups. For practice trials there was little consistency over measurement occasions. For slow trials moderately good

Table 13

Reliabilities of MIT Scores

	Average Practice	Average Slow	Average Practice[1]	Average Fast
Year 1	.27	.67	---	---
Year 2	.29	.69	---	---
Year 3	.35	.63	.45	.60

[1]Scores from practice trials preceding "fast" instructions.

consistency was found. Correcting the correlations for the total longitudinal sample for attenuation yields coefficients of .62, .41, and .49, respectively, for Years 1 x 2, 1 x 3, and 2 x 3. There was no evidence of sex differences in the stability of coefficients, but suggestive evidence for an SES difference, with higher SES children perhaps showing greater consistency over time on slow trials.

Correlations between practice and slow scores averaged around .5 for the three years, with subgroup correlations ranging from .28 to .64 in no apparent pattern. These relations, however, involve subtest-specific correlations; that is, practice and slow times on walking and on drawing each correlated about .5, while practice times from one subtest and slow times from the other correlated around .2. Thus, there is little general response tempo variance reflected in the relation of these two scores.

Table 14

Stability of MIT Scores Over Years for Sex X SES Subgroups

	Low SES			Middle SES			High SES			Total		
	Years			Years			Years			Years		
	1x2	1x3	2x3	1x2	1x3	2x3	1x2	1x3	2x3	1x2	1x3	2x3
Average Practice												
Male	.12	-.02	.18	.05	-.03	.00	.27	.18	.36	.09	.00	.09
Female	.27	-.07	.05	.13	.00	.01	.27	.02	.16	.18	-.03	.05
Total	.19	-.05	.12	.09	-.01	.01	.27	.07	.22	.13	-.01	.07
Average Slow												
Male	.35	.03	.24	.19	.31	.29	.54	.30	.46	.42	.27	.32
Female	.44	.18	.18	.35	.26	.30	.38	.19	.37	.39	.25	.32
Total	.38	.09	.22	.40	.29	.30	.45	.24	.41	.42	.27	.32

A more detailed analysis of score relations from the MIT was made with the Year 3 data in which additional practice and fast administration trials were included. Following the reasoning of Massari, Hayweiser, and Meyer (1969), relations of fast trials to other trial scores can be used to infer the degree to which ability to follow instructions contributes to the child's slow performance. If those children who succeed in slowing their response on request are also those who succeed in speeding their response when asked to do so, it may be that the task is assessing their understanding of instructions rather than their ability to slow their response rate.

Correlational analysis, however, indicated that for each subtest the two practice trials and the fast trials were correlated on the order of .5, while for each subtest the slow administration trial correlated less than .2 with the fast administration. Fast performance was related to practice, not to slow scores.

A principal components factor analysis of the eight time scores obtained in Year 3 emphasizes this point. On Varimax rotation three clear factors emerged. The first was specific to the drawing subtest, and had high loadings (.75 to .84) for the practice and fast trials along with a moderate loading for slow scores (.45). The second presented a comparable picture for the walking subtest--practice and fast trials loading .78 to .80, and slow trial loading .38. The third, finally, had large loadings only for slow trials (.77 and .81), while loadings for the remaining six scores ranged from .31 to -.21. Thus, common variance in slow performance does not reflect either the child's skill in performing the motor activities required or a preferred tempo of response; it is specific to the situation in which the child is asked to perform slowly, i.e., to his self-regulatory ability in this task.

Relations Among Self-Regulatory Behaviors

Table 15 presents intercorrelations of MFF latencies and errors, MIT average practice and average slow trials, MIT average practice (for fast administration) and average fast trials for Year 3 data, and Delay of Gratification scores, separately by year of testing. Similar matrices were obtained for each sex, SES level, and sex by SES level combination; these, however, will not be presented.

The Delay in Gratification task showed no relation to any other measure in the matrix. It should be noted that since this is a one-item task, no

Table 15

Correlations Among Self-Regulatory Scores by Year

	Year 1				Year 2				Year 3				
	B	C	D	G	B	C	D	G	B	C	D	E	F
A MFF Latency	.06	.00	.04	-.01	-.10	.04	.07	.05	-.36	.24	.32	.26	.13
B Error		-.09	-.31	.01		-.08	-.29	.06		-.03	-.26	-.08	.04
C MIT Practice			.51	-.03			.42	.08			.45	.58	.53
D Slow				-.07				.06				.45	.13
E Practice													.60
F Fast													
G Mischel													

reliability index is available for the score. There was, however, no stability of performance over the two years in which the measure was given (r = -.01). Hence, the absence of relations may simply have been due to unreliability of an index depending on a single decision by the child. It may also be noted that in correlational analyses of the larger cognitive battery administered in the first two years of the study, this measure failed to relate to any other performance index. A nonparametric structural analysis of the Year 1 battery did suggest some similarity to performance on a risk-taking index; it may be that factors unrelated to ability to delay gratification, such as the child's trust and confidence in the examiner's promise to deliver the delayed reward, may have been more important than delay ability or preference in this situation.

MIT practice scores had no relation to MFF errors. In the Year 3 data, though not in Years 1 or 2, they did have a low positive relation to MFF latencies (r = .24 and .26 for first and second series of practice trials, respectively). Of greater interest, however, is the relation between MIT slow performance and MFF scores. Here there was an apparent difference across years: at all ages low error scores on the MFF were associated with slow response execution on the MIT; the coefficient was on the order of .3 in each year. MFF latencies, however, were unrelated to MIT slow performance in Years 1 and 2 (r's of .04 and .07), but in Year 3 long latencies were associated with slow MIT performance (r = .32).

Examination of subgroup data indicates that this pattern was found consistently for the various breakdowns of the sample. For the six sex by SES combinations the range of coefficients obtained was as follows: for MFF errors, correlations varied from -.18 to -.47 with a median of

-.26 for males, and from -.11 to - .30 with a median of -.20 for females. For MFF latencies in Years 1 and 2, the coefficients ranged from -.02 to .07 with a median of .05 for males, and from -.15 to .20 with a median of .03 for females. Finally, for MFF latencies in Year 3, they were .41, .32, and .37 for low, middle, and high SES males, and .34, .29, and .19 for females of the three SES levels. No pattern of differences was found with respect to SES. For sex the direction of difference was toward stronger relations for males between MIT slow performance and both MFF errors and latencies. The differences here were small, however, and did not approach significance.

Discriminant Validity with Respect to General Ability

The next question to be considered is that of the relation between self-regulatory behaviors and measures of more general ability or competence. Do the self-regulatory measures provide an assessment of an ability which is distinguishable from the general level of competence at which the child performed in testing? "Distinguishability" does not require that general ability and self-regulatory ability be completely independent of one another, but only that when variance which the self-regulatory measures share with general ability is removed, these measures continue to show positive correlations.

Relations among the three ability measures and their stabilities over years are given in Table 16. General achievement, knowledge of grammar, and understanding of rudimentary quantitative concepts al shared a substantial part of their variance, and did not appear to define separate ability dimensions in the current sample during these early years of development. (In factor analyses of the Year 1 and 2 data, these three measures, along with the remaining cognitive measures in the battery, loaded on a single factor.)

Table 16

Intercorrelations and Stability Coefficients for General Ability Scores

Score			B	C	D	E	F	G	H	I
A	Preschool Inventory	Yr. 1	.75	.78	.43	.54	.55	.36	.60	.49
B		Yr. 2		.83	.39	.67	.59	.36	.69	.56
C		Yr. 3			.46	.62	.64	.37	.65	.63
D	Matched Pictures	Yr. 1				.37	.39	.22	.36	.27
E		Yr. 2					.52	.28	.51	.45
F		Yr. 3						.30	.49	.48
G	Enumeration	Yr. 1							.41	.36
H		Yr. 2								.55
I		Yr. 3								

Table 17 presents the correlations between self-regulatory behaviors and ability measures. As might be expected from their lack of factorial difference, the three ability measures correlated very similarly with each of the self-regulatory behaviors. Where substantial correlations were to be found, the Preschool Inventory generally showed larger correlations than did the other two measures, particularly during the first year of the study. It is likely that this is simply a reflection of the greater reliability of the measure, which is a test of approximately three times the length of either of the other two.

The ability measures were significantly correlated with MFF errors in each year; with MIT slow performance, though not so strongly, in all three years; and with MFF latencies, again not strongly, in only Year 3 data. Thus, there was again evidence that the MFF latencies had a different meaning in Year 3 from that in earlier years. In Years 1 and 2 they bore no relation to measures of competence, either within the self-regulatory domain or in the larger general ability domain; in Year 3 they had moderate but consistent correlations with both kinds of abilities.

Correlation matrices were also obtained for the various sex, SES, and sex by SES subgroups. No differences were found associated with these breakdowns of the data.

Partial correlational analyses were conducted to discover whether significant relations among self-regulatory measures would still be found when variance shared with cognitive indices was removed. The cognitive measure used in these analyses was the Preschool Inventory, which appears to represent the cognitive domain well and which, as a more reliable index, also showed the largest correlations with the self-regulatory behaviors. Table 18 contains correlations among MFF latencies and errors. On the left

Table 17

Correlations Between Self-Regulatory Scores and General Ability Scores by Year

	General Ability Measures		
Self-Regulatory Scores	Matched Pictures	Enumeration	Preschool Inventory
Year 1			
MFF Latency	-.02	-.03	-.03
Errors	-.28	-.22	-.50
MIT Practice	.15	.04	.14
Slow	.29	.20	.46
Mischel	.01	-.00	-.05
Year 2			
MFF Latency	-.02	.01	-.02
Errors	-.49	-.51	-.63
MIT Practice	.11	.09	.11
Slow	.33	.37	.41
Mischel	-.09	-.07	-.09
Year 3			
MFF Latency	.21	.31	.23
Errors	-.43	-.47	-.53
MIT Practice	.05	.08	.05
Slow	.28	.29	.27
Practice	.03	.13	.06
Fast	-.06	-.05	-.07

Table 18

Zero-Order and Partial Correlations Between MFF Scores by Sex and Year

Group		Zero-Order Correlations: Latency X Error	Partial Correlations: Latency X Error (PSI held constant)
Year 1	Male	.09	.07
	Female	.01	.03
Year 2	Male	-.01	-.05
	Female	-.20	-.26
Year 3	Male	-.30	-.24
	Female	-.43	-.34

side of the table are given the zero-order correlations, and on the right are given partial correlations, in which the influence of the Preschool Inventory has been removed. Coefficients are presented separately by sex for each year's data.

Partialling an index of general ability out of the latency-error relation did not reduce this relation substantially for any sex-year combination; thus, the self-regulatory dimension which appears for females in Year 2 and for both sexes in Year 3 is distinguishable, in those subsamples, from general ability. Moreover, there is no support for the argument presented by Lewis et al. (1968) that MFF errors are more a function of the stylistic variable for males than for females, with females' errors more dependent on ability level. In fact, the Year 2 data, showing a significant latency-error relation for females but not for males, and the Year 3 data, showing nonsignificant differences in the direction of stronger latency-error relations for females than for males, indicate a trend in the opposite direction.

Table 19 contains zero-order and partial correlations between MIT slow administration scores and the MFF latencies and errors; again the Preschool Inventory score was the score partialled out. For both sexes, removing variance shared with a general ability reduced the relation of MFF errors and MIT slow times to nonsignificance; thus, their relation appears to have been based on the portion of each that represents general level of competence. The correlations of latencies with the MIT slow score, however, present a different picture. In Years 1 and 2, where the zero-order correlation was near zero, the partial correlation was also. In Year 3 data, however, there was evidence that MFF latencies shared

Table 19

Zero-Order and Partial Correlations Among MFF and MIT Scores by Sex and Year

Group		Zero-Order Correlations		Partial Correlations	
		MFF Errors X MIT Slow	MFF Latency X MIT Slow	MFF Errors X MIT Slow (PSI held constant)	MFF Latency X MIT Slow (PSI held constant)
Year 1	Male	-.31	.04	-.11	.08
	Female	-.30	.05	-.09	.04
Year 2	Male	-.29	.06	.00	.09
	Female	-.27	.08	-.08	.10
Year 3	Male	-.27	.33	-.12	.29
	Female	-.23	.30	-.04	.21

variance with the MIT slow performance which was not general ability variance. Thus, it appears that, along with a developing relevance for performance of the reflective-impulsive style, as measured by the MFF, there is a developing relation between stylistic and ability measures of self-regulation. The partial correlations are relatively low but reach a satisfactory level of confidence (for males, $r = .29$, $p < .0001$; for females, $r = .21$, $p < .0001$); these coefficients do not differ significantly from one another ($p < .25$).

Discriminant Validity with Respect to Response Tempo

It is apparent in the data examined thus far that the meaning of the self-regulatory measures showed a change over age, toward increasing convergence of the several measures in this domain. It is also apparent that the common variance among these measures is distinguishable from general ability variance. It remains to be seen, however, to what degree these measures provide an index of a more general dimension having to do with the child's preferred speed of response, irrespective of the cognitive demands of the task. The only evidence which has been presented with respect to this issue involves the correlational differentiation between MIT slow performance and the practice and fast administration scores on MIT; this differentiation argues that slow performance, and the MFF scores to the degree that they share variance with MIT slow performance, depend on something other than such a response tempo dimension. Here, the several additional time scores will be examined to provide further evidence on response tempo.

One indication that response tempo is distinct from self-regulatory ability is given by an examination of mean changes for tempo scores over

Table 20

Means for Response Tempo Scores by Year

Score	Year 1		Year 2		Year 3	
	Mean	SD	Mean	SD	Mean	SD
Seguin Form Board-Fastest time.[1]	1.63	0.20	1.44	0.13	----	----
Preschool Embedded Figures Test-Latency to first response.[2]	.842	.174	.752	.169	.756	.136
Sigel Object Categorization Test-Latency to first choice.[2]	.799	.216	.764	.202	.826	.179

[1] In minutes

[2] Transformed by log (X).

years. Table 20 provides means for these measures. Time taken on fastest trial on the Seguin Form Board, which was given only in the first two years of data collection, decreased from the first to the second year by about one standard deviation. Latency to first response on the Preschool Embedded Figures Test also decreased from the first to the second year, and did not change from the second to the third year. Here, both the linear and the quadratic age trend were significant ($p < .0001$). Finally, latency to first object choice on the Sigel Object Categorization Test showed a significant quadratic trend ($p < .0001$); latencies decreased from the first to the second year, and then increased to the third year. None of these measures, then, followed the pattern of monotonic increases with age which were found for MFF latencies, for MFF accuracy (here, a decrease in error scores), and for MIT slow performance.

Correlations over time and across measures for the tempo scores showed the Seguin score to differ in several respects from the Sigel task and the PEFT. Seguin fastest time had high stability from the first to the second year's testing ($r = .53$), while the other two scores had low stability coefficients, ranging from -.03 to .20. Also, within year the Sigel and PEFT scores had a low but quite consistent positive correlation (ranging from .20 to .23 for the three years). The Seguin scores also correlated positively with the two latency scores, but with a range from .01 to .17, and a median correlation of .07. The same patterns of correlations were found when data were examined for the various sex and SES subgroups.

Correlations between these three measures and the three general ability measures under examination in this report also show differences. The Seguin score correlated negatively with the ability measures; that is, fast times on Seguin were associated with higher scores on the Preschool Inventory, Matched Pictures, and Enumeration tests. Time on this task thus appears to index perceptual-motor coordination, ability to follow instructions, or other components of ability represented in the general ability domain. Sigel and PEFT latencies, on the other hand, had no consistent correlation with the ability measures; over the three years for which correlations were available, the range of coefficients was from -.19 to .10, with a median of -.03.

Table 21 presents the correlations between self-regulatory behaviors and the Sigel and PEFT latencies, for the total longitudinal sample. Response tempo measures had negligible correlations with MFF errors, with MIT practice and slow scores, and with the Delay of Gratification measure. Sigel latencies, however, were moderately correlated with MFF latencies in Year 1, had a negligible correlation with them in Year 2, and showed a moderate positive correlation in Year 3. PEFT latencies had a lower

Table 21

Correlations of Response Tempo Scores
With Self-Regulatory Scores by Year

Self-Regulatory Scores	Response Tempo Scores					
	Year 1		Year 2		Year 3	
	Sigel RT	PEFT RT	Sigel RT	PEFT RT	Sigel RT	PEFT RT
MFF Latency	.46	18	.07	-.10	.29	.18
Errors	.11	-.11	-.03	.06	-.02	.07
MIT Practice	.02	.03	.08	-.07	.18	.07
Slow	.02	.12	.09	.04	.11	.07
Practice	---	---	---	---	.19	.08
Fast	---	---	---	---	.15	.03
Mischel	.00	.03	.03	.06	---	---

correlation with MFF latencies but followed a similar pattern, in that the Year 1 and 3 correlations were higher than that in Year 2.

This pattern was repeated throughout the various sex and SES subgroups; a detailed presentation of the correlations involving MFF latencies is given in Table 22.

Table 22

Correlations of Response Tempo Scores With MFF Latencies for Sex X SES Subgroups by Year

Group	Year 1 Sigel RT	Year 1 PEFT RT	Year 2 Sigel RT	Year 2 PEFT RT	Year 3 Sigel RT	Year 3 PEFT RT
All Ss	.46	.18	.07	.10	.29	.18
Male	.43	.16	.05	.13	.31	.25
Female	.49	.21	.07	.07	.28	.11
Low SES	.46	.25	-.02	.18	.20	.20
Middle SES	.47	.16	.08	.09	.32	.20
High SES	.44	.23	.25	.18	.37	.17
Male Low SES	.50	.28	.04	.28	.10	.21
Middle SES	.41	.10	.10	.11	.36	.25
High SES	.36	.20	.12	.34	.41	.31
Female Low SES	.42	.21	-.12	.14	.30	.19
Middle SES	.55	.23	.06	.11	.28	.14
High SES	.46	.22	.35	.00	.38	.10

These data suggest that the MFF latencies, during the first year of the study, were largely indexing response tempo. By the second year this was no longer true; their relations with tempo measures had dropped to near zero. In the third year, however, moderate positive relations with the tempo measures again appeared. Perhaps at this point in development, Sigel latencies may also have begun to show differences associated with a reflective or impulsive style. Recalling the earlier discussion of the Sigel test, if the child is faced with uncertainty as to which he will perform of several groupings he is able to make the conditions for elicitation of stylistic variance are present; around age six some of the children in the sample may indeed have become capable of seeing and reflecting upon several alternative grouping possibilities.

Discussion

Several aspects of the results presented here converge to indicate a change with age in the meaning of the self-regulatory behaviors; all involve changes in the correlations of MFF latencies with other scores. To summarize, MFF latencies, from the pre-nursery school to the kindergarten year, go 1) from no relation to MFF errors, to a significant negative relation; 2) from no relation to MIT slow scores, to a low but significant positive relation; and 3) from a moderately strong positive relation to latencies on the Sigel Object Categorization Test, to no relation, and then to a moderate positive relation. The latencies also go from no relation to Preschool Inventory scores, representing the general ability domain, to a low positive correlation in Year 3.

The results from the Year 3 testing are consistent with thos which have been found by other investigators working in this domain. Kagan and his collaborators have shown in many studies with elementary school children the negative latency-error relation which appeared here (e.g., Kagan, Pearson, and Welch, 1966; Kagan & Kogan, 1970), and have found a positive relation between MFF latencies and the ability to perform a motor act slowly (Kagan et al., 1964; see also Hess, Shipman, Brophy & Bear, 1969; Ward, 1968b). The latency-error relation has been seen in children below the elementary school age (Harrison & Nadelman, 1972; Katz, 1971; Lewis et al., 1968; Meichenbaum & Goodman, 1969; Mumbauer & Miller, 1970; Ward, 1968a; Wright, 1972; Zucker & Stricker, 1968) as has a positive relation between MIT slow performance and MFF latency (Harrison & Nadelman, 1972). Banta (1970) also reported a positive relation between MFF and MIT performance; but his MFF score was based solely on errors and so may have related to MIT only through variance shared with general ability.

The present results differ from those in all the studies reported above in that there was no relation between MFF latencies and errors in the first two years. There is, however, evidence that the difference is one of populations rather than of contradictory findings. Age and SES characteristics of the children studied in the eight investigations with preschool children which are listed above are given in Table 23. Note that in five of the eight studies, the samples were exclusively middle class, and that in four of them, the children were of kindergarten age--actually in kindergarten classes in two studies, and with an average age ranging from 61 to 69 months in all four studies. Further, there are no samples in this list in which the children were both of lower SES and of an average age less than five years. It is easily assumed, from the number of reports which have provided positive findings with respect to the MFF in both elementary and preschool samples, that the negative relation between latencies and errors has been found over the entire preschool age range in populations including all levels of SES; but in fact the low SES--younger age range, to which the present sample corresponds, appears not to have been examined in these studies.

These considerations suggest that for many children the ages from around three to five represent a transition period in the organization of self-regulatory behaviors. Children who are very bright or who are highly advantaged may complete this transition earlier, perhaps by around the age of three to four. Wright's (1972) youngest subsample, ranging in age from 34 to 49 months, and Lewis et al.'s (1968), with 44 month olds, both showed significant negative relations between latencies and errors, implying that these children had reached a point in development at which long latencies were employed in the service of superior information processing.

Table 23

Age and SES Characteristics of Preschool Samples in MFF Studies

Reference	Age	SES
Harrison & Nadelman, 1972	48-60 months	Middle Class
Katz, 1971	44-65 months	(Private Preschool)
Lewis et al., 1968	44 months	Middle Class
Meichenbaum & Goodman, 1969	63-76 months	Middle Class
Mumbauer & Miller, 1970	56-68 months	Advantaged & Disadvantaged
Ward, 1968a	50-78 months	Middle Class
Wright, 1972	34-80 months	Middle Class
Zucker & Stricker, 1968	57-65 months	Middle Class & H[illegible]. start

Children who are not so favored, however, appear to reach this point at a later age. The present sample as a whole had not begun this transition at the average age of 3 1/2, i.e., in the first year of testing; it had completed it by the age of 5 1/2, in the third year of testing. Thus, the question is one of rate of development, not of a different developmental course, except to the degree that a longer period of less adequate self-regulation has implications for the child's learning during this time. Note, also, that the difference suggested is one of the integration of behaviors, not of the mean levels of performance. In the present sample all children were responding very quickly relative to the time taken by older children on the MFF, and relative as well to the time they would have needed in order to perform the task in a more or less error-free manner. The transition which is proposed as occuring during the 3-5 year old range, finally, should not be confused with that which White (1965) and others have suggested comes around the time of entry into the elementary school system, and which may be in part dependent on the additional cognitive requirements the formal educational system imposes on the child's performance.

This interpretation is post hoc; however, it is given additional support by a consideration of results obtained when the present data were analyzed by the sex and SES of the child. With respect to SES, Year 2 latency-error correlations were .03 for the lowest SES group, -.17 for the middle group, and -.30 for the high group. Thus, the least advantaged children showed no signs of development of the expected relation during this year; the most advantaged showed performance comparable to that obtained one year later by the sample as a whole; and the middle SES group showed intermediate results. These data clearly fit the hypothesis of a differential rate of growth in the relevance of the stylistic variable for performance. Likewise, in looking at results according to the sex of the child, in Year 2 males showed a latency-error correlation of -.01, while females showed one of -.20. If it is assumed that females are generally more cognitively advanced during the preschool years, these coefficients also fit the differential growth hypothesis. The weight of such evidence as there is favors this assumption (Maccoby, 1966). In the present sample evidence for female superiority in performance on ability measures was mixed. Two of the three ability scores examined for this report showed a significant sex difference favoring females ($p < .001$), while the Preschool Inventory performance exhibited a nonsignificant trend in the same direction in each of the three years of testing. On the other hand, analyses based on performance of the larger sample from which this one was drawn generally failed to show sex differences in the first two years. The absence of a sex by year interaction suggests that this difference is not a function of changes over age in sex differences; perhaps it reflects some contingency among sex, ability, and the child's continued availability over time for testing.

The question of sex differences in self-regulatory behaviors is an important one. Maccoby (1966) has summarized evidence suggesting that the meaning of an impulsive personal style varies with sex. Intellectual functioning appears to be associated positively with indices of impulse expression in females, and negatively with the same variables in males. If males and females generally differ, such that most females are in some sense too highly controlled, and most males too little, these findings indicate that a less extreme position with respect to impulse expression is optimal. The existing data with respect to the more cognitive self-regulatory behaviors examined in this report, however, do not conform to this generalization. Rather, where relations have been found, they indicate that for both sexes general intellectual functioning is higher in those who are more controlled. The present report is not addressed to these issues, except in so far as it examines general ability correlates of the self-regulatory measures employed; and here, where there were relations between self-regulation and general ability, they were uniformly positive. Subsequent reports, in which implications for a broader range of cognitive and noncognitive behaviors are examined, should contribute more fundamental information.

It should be emphasized that the younger children in the present sample were not in any sense totally lacking in self-regulatory ability. On the MIT, for example, even the youngest children tested in the first year of the study were able to slow their response execution upon request. Moreover, while they were unable to withhold response on MFF trials long enough to insure high accuracy, they responded only a little more quickly (and with very similar error rates) than the middle-class three-year-olds tested by

Lewis et al. (1968), where at least in males the reflection-impulsivity dimension was clearly manifested. It is the organization of these behaviors, not the level, which is at issue.

Perhaps these results can be interpreted as another instance of the developmental sequence proposed by Flavell and others, which was alluded to in the introduction. The young child is not lacking in self-regulatory ability, as seen in his MIT performance. What he lacks is an understanding that it is to his advantage to bring this ability to bear in his test-taking performance. With increasing age and experience, he may simply learn that abilities he has and uses on command are relevant for task performance, and so will come to apply them without specific instruction.

A related possibility is that the young child's goals in the task may not be those of the tester; for some children, pleasing the examiner, being "fair" by making sure that their choices are distributed over all the options, and similar "task irrelevant" objectives may be equally as important as obtaining high accuracy. Part of the set with which the older or more advantaged child approaches the task may include a greater orientation toward intellectual achievement; and so the developmental change may reflect not just changes in recognizing the relevance of an ability for task performance, but also changes in an understanding of what an adult is requesting in setting the task.

Kagan et al. (1964) suggested an alternative explanation of the origins of self-regulatory ability, proposing that cognitive impulsivity might be an outgrowth of a general, perhaps physiologically based, quick response tempo. From this perspective it could be argued that the ability to withhold or delay response would increase as part of a general developmental sequence;

and that only when the child has reached a point at which a moderate delay is possible can he bring his cognitive abilities fully to bear on task performance. The present data are less compatible with this possibility than with the cognitive interpretation in several respects. First, as has been indicated, all the children tested showed some ability to slow their response rate on command. It is not clear why children should be physiologically unable to delay on the MFF but able to do so on the MIT. Second, there was little evidence for the general response tempo dimension which this explanation implies should be found in young children. MIT practice scores showed little stability across subtests or across years, and little relation to other response tempo indices examined. These indices themselves had low intercorrelations and stabilities. The only notable exception to this pattern was the moderate relation between MFF and Sigel latencies in Year 1. Third, this argument implies that indices of response tempo in the early years should be predictive of self-regulatory behaviors at a later point in development. The relevant correlations (not presented) did not show this relation.

The present report, in discussing correlational changes with age, involves inferences from group data to changes over time in the task approach taken by individuals. Some questions as to individual changes remain unanswered through this approach. For example, given that some individuals became "reflective" (slow and accurate) on the MFF in Year 3, were these more likely in Years 1 and 2 to have been "impulsive" (fast and inaccurate), fast and accurate, or slow and inaccurate? The product moment correlations examined indicate only that there was little overall consistency across years in latency considered alone; it remains possible that there is some predictability in movement from one of these subgroups to another. As another example, once individuals can be described as falling at some point on the reflective-impulsive dimension, what is the course of their change from year to year? Does the equation expressing the regression of latencies on errors remain constant over time, with individuals changing in both latencies and errors in such a way as to fall on the same regression line; or does the regression equation change systematically, perhaps with the same latency at a later age being associated with a lower error score? The data available, especially when combined with data from these tasks obtained in later years of the study, will permit examination of some of these questions. It is possible that sequences of individual change will be as important as information on overall level of performance for understanding the ways in which family and educational variables affect children's development.

Finally, though the increasing integration of self-regulatory behaviors with age has been the focus of attention, the six-year-olds in the present study still showed a relatively low level of such integration. Approximately 12% of the variance in MFF latencies was shared with MFF errors, and only 4 to 8% with MIT slow scores. Correction for unreliability in these scores would raise the percentages somewhat, but would not change the general level of relation dramatically: thus, in a sense the various self-regulatory scores are much more independent of one another than they are related. One of the concerns in further analyses of the data will be to discover whether children whose MFF latencies and MIT slow scores are discrepant--those with long latencies but fast times on slow trials, and those showing the converse pattern--are intermediate in their cognitive and personal characteristics to those whose performance shows greater consistency, or whether they are divergent from expectations based on an average of these two instances of self-regulatory opportunities.

References

Banta, T. J. Tests for the evaluation of early childhood education: The Cincinnati Autonomy Test Battery (CATB). New York: Brunner/Mazel, 1970.

Corsini, D., Pick, A., & Flavell, J. Production deficiency of nonverbal mediators in young children, Child Development, 1968, 39, 53-58.

Eska, B., & Black, K. N. Conceptual tempo in young children. Child Development, 1971, 42, 505-516.

Fraser, C., Bellugi, U., & Brown, R. Control of grammar in imitation, comprehension, and production. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 121-135.

Harrison, A., & Nadelman, L. Conceptual tempo and inhibition of movement in black preschool children. Child Development, 1972, 43, 657-668.

Hess, R. D., Shipman, V. C., Brophy, J. E., & Bear, R. M. The cognitive environments of urban preschool children: Follow-up phase. Chicago: University of Chicago Press, 1969.

Kagan, J. Developmental studies of reflection and analysis. In A. H. Kidd & J. L. Rivoire (Eds.), Perceptual Development in Children. New York: International Universities Press, 1966. Pp. 487-522.

Kagan, J., & Kogan, N. Individuality and cognitive performance. In P. H. Mussen (Ed.), Carmichael's Manual of Child Psychology. Volume 1. New York: John Wiley & Sons, 1970.

Kagan, J., Pearson, L., & Welch, L. Conceptual impulsivity and inductive reasoning. Child Development, 1966, 37, 583-594.

Kagan, J., Rosman, B. L., Day, D., Albert, J., & Phillips, W. Information processing in the child: Significance of analytic and reflective attitudes. Psychological Monographs, 1964, 78, (1, Whole No. 578)

Katz, J. M. Reflection-impulsivity and color-form sorting. Child Development, 1971, 42, 745-754.

Lewis, M., Rausch, M., Goldberg, S., and Dodd, C. Error, response time and IQ: Sex differences in the cognitive style of preschool children. Perceptual and Motor Skills, 1968, 26, 563-568.

Luria, A. Verbal regulation of behavior. In M. A. B. Brazier (Ed.), The Central Nervous System and Behavior. (Report of the third Macy Conference.) Madison, N. J.: Macy Foundation, 1966.

Maccoby, E. E. Sex differences in intellectual functioning. In E. E. Maccoby (Ed.), The Development of Sex Differences Stanford, Calif.: Stanford University Press, 1966.

Maccoby, E. E., Dowley, E. M., Hagen, J. W., & Degerman, R. Activity level and intellectual functioning in normal preschool children. Child Development, 1965, 36, 761-770.

Massari, D., Hayweiser, L., & Meyer, W. Activity level and intellectual functioning in deprived preschool children, Developmental Psychology, 1969, 1, 286-290.

Meichenbaum, D., & Goodman, J. Reflection-impulsivity and verbal control of motor behavior. Child Development, 1969, 40, 785-797.

Meissner, J. A., & Shipman, V. C. Disadvantaged children and their first school experiences: Summary Report: Structural stability and change in the test performance of urban preschool children. Princeton, N. J.: Educational Testing Service, 1973, PR-73-8. Prepared under Grant H-8256, Department of Health, Education, and Welfare.

Mischel, W. Preference for delayed reinforcement: An experimental study of a cultural observation, Journal of Abnormal and Social Psychology, 1958, 66, 57-61.

Moely, B., Olson, F., halwes, T., & Flavell, J. Production deficiency in young children's clustered recall, Developmental Psychology, 1969, 1, 26-34.

Mumbauer, C. C., & Miller, J. O. Socioeconomic background and cognitive functioning in preschool children. Child Development, 1970, 41, 471-480.

Shipman, V. C. Disadvantaged children and their first school experiences: Structure and development of cognitive competencies and styles prior to school entry. Princeton, N. J.: Educational Testing Service, 1971, PR-71-19. Prepared under Grant H-8256, Department of Health, Education, and Welfare.

Shipman, V. C. Disadvantaged children and their first school experiences: Demographic indexes of socioeconomic status and maternal behaviors and attitudes. Princeton, N. J.: Educational Testing Service, 1972a, PR-72-13. Prepared under Grant H-8256, Department of Health, Education and Welfare.

Shipman, V. C. Disadvantaged children and their first school experiences: Structural stability and change in the test performance of urban preschool children. Princeton, N. J.: Educational Testing Service, 1972b, PR-72-18. Prepared under Grant H-8256, Department of Health, Education, and Welfare.

Shipman, V. C. Disadvantaged children and their first school experiences: Technical report series. Princeton, N. J.: Educational Testing Service, 1972c, PR-72-27. Prepared under Grant H-8256, Department of Health, Education, and Welfare.

Ward, W. C. Creativity in young children. Child Development, 1968a, 39, 737-754.

Ward, W. C. Reflection-impulsivity in kindergarten children. Child Development, 1968b, 39, 867-874

White, S. Evidence for a hierarchical arrangement of learning processes. In L. Lipsitt and C. Spiker (Eds.), Advances in child development and behavior (Vol. 2). New York: Academic Press, 1965.

Wozniak, R. Speech-for-self as a multiple reafferent source of nonspeech motor control. Paper presented at the meeting of the Society for Research in Child Development, Philadelphia, April 1973.

Wright, J. C. The KRISP: A technical report. In J. C. Wright, Attention and cognitive style. Kansas Center for Research in Early Childhood Education, 1972.

Zucker, J. S., & Stricker, G. Impulsivity-reflectivity in preschool Head Start and middle class children. Journal of Learning Disorders, 1968, 1, 578-584.