DOCUMENT RESUME

ED 116 777 95 PS 008 196

AUTHOR Stephens, Mark W.
TITLE Effects of Type of Preschool Experience and Socioeconomic Class on Academic Achievement Motivation. Final Report.
INSTITUTION Purdue Research Foundation, Lafayette, Ind.
SPONS AGENCY National Center for Educational Research and Development (DHEW/OE), Washington, D.C. Regional Research Program.
BUREAU NO BR-0-E-117
PUB DATE Aug 73
GRANT OEG-5-70-0039
NOTE 38p.

EDRS PRICE MF-$0.76 HC-$1.95 Plus Postage
DESCRIPTORS Comparative Analysis; Cultural Factors; Economic Disadvantagement; Intervention; *Locus of Control; *Measurement Instruments; *Motivation; Observational Learning; *Preschool Education; Preschool Programs; Question Answer Interviews; Social Reinforcement; *Socioeconomic Influences; Student Motivation; Test Reliability; Validity
IDENTIFIERS Contingency Interview; *Experimenter Characteristics; Project Head Start; Stephens Delys Reinforcement

ABSTRACT

Four experiments were designed to identify socioeconomic differences in preschool locus of control, develop a measurement technique for differentiating between internal and external locus of control in preschoolers, and study the effect of four kinds of preschool programs on locus of control. During the first experiment, the Stephens-Delys Reinforcement Contingency Interview (SDRCI) was developed to assess internal locus of control development in preschoolers. When used with 24 4-year-olds in a Head Start program, the measure was found to have rater and retest reliability; the race of the interviewer did not significantly affect scores. The second experiment indicated that the performance of 32 preschool boys on a mirror-tracing task was positively related to internal locus of control as measured by the SDRCI. In the third study, investigators tested 55 Head Start preschoolers and 50 middle-class nursery school children with the SDRCI. Lower internal control scores were found for the Head Start children than for the middle-class nursery school group; no differences were found between black and white Head Start groups. A final study of 114 children found a nonsignificant tendency for Montessori preschool experience (and to a lesser extent, parent cooperative nursery school experience) to increase internal control, as measured by the SDRCI, more than Head Start or a more structured compensatory preschool program. (Author/BRT)

PS

Final Report

Project No. 0-E-117
Grant No. OEG-5-70-0039

Mark W. Stephens, Ph.D.
Purdue Research Foundation
School of Humanities, Social Science, and Education
Department of Psychological Sciences
Purdue University
West Lafayette, Indiana 47907

EFFECTS OF TYPE OF PRESCHOOL EXPERIENCE AND SOCIOECONOMIC CLASS ON ACADEMIC ACHIEVEMENT MOTIVATION

August 1973

U. S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Office of Education

National Center for Educational Research and Development

(Regional Research Program)

Project No. 0-E-117
Grant No. OEG-5-70-0039

EFFECTS OF TYPE OF PRESCHOOL EXPERIENCE AND SOCIOECONOMIC CLASS ON ACADEMIC ACHIEVEMENT MOTIVATION

Mark W. Stephens, Principal Investigator

Abstract

This research developed the Stephens-Delys Reinforcement Contingency Interview to assess internal control expectancy development and thus investigate the need for redesigning compensatory preschool programs to expedite development of internal control expectancies. Four studies were performed. The first, with 24 Head Start children, found adequate rater and retest reliability and freedom from interviewer ethnicity effects on scores. The second, with 32 nondisadvantaged day care center preschool boys, found a significant relation of internal control scores to performance on a mirror-tracing task, thus supporting construct validity of the measure. The third, with 55 Head Start and 50 nondisadvantaged nursery schoolers, found lower internal control scores in the former than the latter, with no difference between black and white Head Start groups. The final study, with 114 children, found a nonsignificant tendency for Montessori preschool experience (and, to a lesser extent, parent cooperative nursery school experience) to increase internal control expectancies more than either Head Start or a more structured compensatory preschool program. It was concluded that more extensive research is needed, but compensatory preschool programs may need to be redesigned to better enhance internal control expectancy development.

Final Report

Project No. 0-E-117

Grant No. OEG-5-70-0039

EFFECTS OF TYPE OF PRESCHOOL EXPERIENCE AND SOCIOECONOMIC CLASS ON ACADEMIC ACHIEVEMENT MOTIVATION

Mark W. Stephens
Purdue Research Foundation
West Lafayette, Indiana 47907

August 1973

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

U.S. DEPARTMENT OF

HEALTH, EDUCATION, AND WELFARE

Office of Education

National Center for Educational Research and Development

TABLE OF CONTENTS

Table of Contents (Continued)

LIST OF TABLES

Project No. 0-E-117
Grant No. OEG-5-70-0039

EFFECTS OF TYPE OF PRESCHOOL EXPERIENCE AND SOCIOECONOMIC CLASS ON ACADEMIC ACHIEVEMENT MOTIVATION

Mark W. Stephens, Principal Investigator

INTRODUCTION

Most compensatory preschool education programs for the culturally disadvantaged have been aimed primarily at cognitive training. They have provided enrichment, stimulation, and specific training to increase language, prereading skills, pre-arithmetic concepts, number recognition, and other cognitive skills. Yet one of the major conclusions of the classical "Coleman report" on Equality of Educational Opportunity (Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld, & York, 1966) was that "... a stronger relationship to achievement than...all the 'school' factors together, is the extent to which an individual feels that he has some control over his own destiny," and further that "... minority pupils, except for Orientals, have far less conviction than whites that they can affect their own environments and futures. When they do, however, their achievement is higher than that of whites who lack that conviction (p. 23)." This "conviction" is obviously a motivational variable, not a cognitive skill susceptible to development as a simple function of enrichment, stimulation, and readiness training. Nevertheless, little attention has been directed toward using compensatory preschool experiences as a means of forestalling this motivational handicap; and, in fact, some aspects of programs emphasizing cognitive training may even reinforce inclinations toward a passive, fatalistic attitude toward educational accomplishment. This project was concerned with the general problem of whether compensatory programs can and should be designed to combat this motivational problem. The specific aims were to investigate (1) whether there are socioeconomic differences in this variable ("the conviction they can affect their own environments and futures") among children of preschool age, (2) what is the effect of different kinds of preschool programs on this variable.

The variable to which the Coleman report alludes is virtually identical to the variable Rotter (1966) has called "locus of control," abbreviated "IE" (for "internal" vs. "external" control of reinforcement). In Rotter's theory, IE is a higher-order expectancy variable: "internal" denotes the generalized expectancy that reinforcement is contingent upon or controlled by one's own behavior, as in an operant conditioning paradigm, and "external" denotes the expectancy that reinforcement is controlled by external forces (including "luck") or other people and not contingent on one's own behavior, as in a respondent conditioning paradigm. IE research has been largely concerned with the antecedents and consequences of individual differences in such expectancies; social psychological research on causal attribution (e.g., Kelley, 1973; DeCharms, 1968) has approached many of the same phenomena with interest primarily in the situational determinants of such expectancies, or "attributions." Similar phenomena have been investigated by Solomon and others in their research on "conditioned helplessness" in dogs (see Maier, Seligman, & Solomon, 1969); these authors conceptualize such phenomena in S-R learning theory terminology, pertinent to mediation of instrumental learning. Both Logan's (1960) and

Spence's (1960) S-R treatments of instrumental learning touched on the contingency of the "anticipation" of reinforcement to specific instrumental responses, as did McClelland's (McClelland, Atkinson, Clark, & Lowell, 1953) and Atkinson's (1964) motivation theories. Rotter's IE variable, and associated research and theory, was used as the primary point of departure for this project because it offered the greatest body of directly relevant data and methodology to draw from.

Several independent studies have confirmed Coleman's finding of socioeconomic differences in IE (e.g., Battle & Rotter, 1963; Franklin, 1963; Nowicki & Strickland, 1973); it is consistently found that lower class blacks have significantly more external expectancies than either middle class blacks or lower or middle class whites, with the latter three groups differing little, if at all, from one another. Other studies have confirmed the role of IE expectancies in mediating socioeconomic differences in coping with behavior (e.g., Lefcourt, 1965). There are also several studies which confirm the role of IE in academic achievement, both among college students (Stephens, 1969; Butterfield, 1964) and among elementary school students (Crandall, Katkovsky, & Preston, 1962). What has not been known, however, is (1) whether there are discriminable IE differences related to socioeconomic status as early as preschool age; (2) whether IE differences at this age are related to acquisition of cognitive skills; (3) whether different kinds of preschool experiences do or can affect IE expectancies regarding learning in school; or (4) whether techniques might be developed which could specifically enhance development of appropriate internal control expectancies regarding school achievement.

Prior to this project these questions could not be answered because no IE measure for preschool children had yet been developed. At least six IE measures had been devised for elementary school children (Bialer, 1961; Battle & Rotter, 1963; Crandall, Katkovsky, & Crandall, 1965; Shore, Milgram, & Malasky, 1971; Gruen, 1970; Nowicki & Strickland, 1973). None, however, could be used with children of preschool age. Consequently, for the present research it was necessary first to devise and to assess the reliability and validity of an IE measure which could be used with preschool age children.

This project entailed four relatively distinct phases, each reported separately below: (1) the development and initial assessment of reliability and convergent validity of an IE measure that could be used with preschool children; (2) construct validation of the measure against a nonverbal performance criterion; (3) assessment of differences between disadvantaged and nondisadvantaged preschool age children on this measure; and (4) initial assessment of the impact on IE, using this measure, of at least three different kinds of preschool programs (Head Start, parent cooperative, and Montessori nursery schools).

STUDY I: DEVELOPMENT OF AN IE TEST FOR PRESCHOOL CHILDREN

Introduction

Pilot testing confirmed that the previously developed IE tests for children are all beyond the cognitive limits of preschool age subjects. For example, the Intellectual Achievement Responsibility (IAR) test (Crandall, et al., 1965) is the test which had been most used in children's IE research;

but Crandall (personal communication) does not advise using it with children below about third grade level. It is a forced-choice test, and must be orally presented for children of limited reading skill. Crandall reports that subjects below third grade level show a significant tendency to repeat the last-read response alternative, presumably because of difficulty in remembering the first-read alternative. Pilot testing for this project confirmed this phenomenon. Gruen (1970) developed a test designed to avoid or minimize this problem. He used pictures, representing the response alternatives, to supplement the oral presentation of response alternatives and thereby reduce the child's difficulty in remembering the first-read response. This test, however, was designed for children of second grade level and above and (like the IAR) is specific to school achievement situations and therefore inappropriate for preschool children. The Nowicki-Strickland IE (NSIE) test (Nowicki & Strickland, 1973) is a yes-no format test and is not specific to school achievement situations; but it is not recommended for use with children below second grade, and Stephens (1971a, 1972b) found a Kuder-Richardson 20 reliability estimate of only 0.32 for the NSIE among 575 second graders. Here again, pilot testing for the present study found strong response set tendencies among preschoolers (tendencies to repeat "yes" or "no" responses) apparently reflecting the child's inability to understand the test questions. Tests in the conventional forced-choice or yes-no formats seemed, then, not to be feasible for children of preschool age even when they deal with situations familiar to the child and even when they present no problem of remembering response alternatives. It was apparent that something other than a yes-no or forced-choice test would be needed.

Pilot work for the present study revealed another problem: that concepts such as "luck" and "skill" were incomprehensible to most nursery schoolers tested, and even less abstract terms were incomprehensible to many. It was obvious, then, that in order to test preschoolers it would be necessary to make the test questions as concrete and cognitively simple as possible. Pilot testing also revealed that it was far more difficult for the child to give even a yes-no answer to a meaningful IE proposition than to make his own verbal production to a very simple open-ended question, as in free-response test procedure. This suggested that some free-response method, with maximally simple questions, would be necessary to elicit meaningful responses from preschoolers.

Since conventional forced-choice or yes-no test procedures could not be used with preschoolers, it was necessary to refer to the theoretical model in which IE is defined (Rotter, 1966) to see what alternative measurement models might be possible. As Jessor and Hammond (1957) pointed out, a measure of a theoretical construct should, ideally, correspond in form to the theoretical definition of that construct; that is, the behavior observed should be behavior which, in that theory, is denotatively representative of that construct. Since IE is, in Rotter's model, an expectancy variable, ideally an IE measure might be a series of statements of expectancy (preferably in a zero to 100 percent, rather than categorial present-absent, form). IE is actually an expectancy about expectancies. It is the subject's expectancy that his own behavior would change the probability that reinforcement might occur. For example, a child may have an expectancy of 0.10, or 0.85, or anything between 0.00 and 1.00 that his teacher will be approving of him whether or not he does anything to elicit approval; regardless of his "confidence" that his teacher will be approving, his IE expectancy is how much he thinks this probability

will change (or how probable he thinks it is that this probability will change at all) if his behavior changes. The theoretically purest IE question might take this form: "I have a ___ percent chance of teacher liking me if I try to make her like me and a ___ percent chance if I don't," the IE expectancy being the difference between these two probability statements; or, alternatively, "There is a ___ percent chance that the probability of the teacher's liking me will change if I try to make her like me."

Obviously, it is impossible to elicit a quantified probability statement from a preschool-age child. However, it is possible to quantify the degree of association he shows between reinforcing events on the one hand and his own behavior on the other. The probability of a child associating the occurrence of a reinforcement with his own behavior, rather than with some other contingency, can be assumed to reflect his perception of the probability that his behavior is the most salient contingency for that reinforcement. An index of associative strength might be, in fact, a more direct index of subjective probability of relation of behavior to reinforcement than would response choice on a more transparent true-false or force-choice test, which is vulnerable to influence by the subject's expectancies regarding which given response alternative may be considered the more socially desirable. A free-response measure, then, might give a more straight-forward index of associative strength and, therefore, of perceived relationship or subjective probability of relationship of reinforcement to behavior than could any kind of limited-response-choice measure. Since preschool-age children must, of course, be tested individually and orally anyway, a free-response measure would not be likely to be any more time-consuming than a conventional test format.

The general model, then, called for a way of determining, by a free-response method employing maximally simple questions, the degree of association between reinforcements and behaviors. Presumably one might use either of two types of questions to assess this strength of association. One type would posit the occurrence of a behavior and determine whether or not the child associates with that behavior some reinforcement or something else: for example, "What happens when you listen closely to the teacher?" This kind of question was tried in pilot testing; but there seemed to be too much obvious stimulus pull, for any behavior that might generally be expected to be subject to reinforcement, to yield individual differences in responses. The other question type would posit the occurrence of some reinforcement and ask what, in effect, are the contingencies for occurrence of such an event--and then observe whether the child cites some behavior of his own or cites someone else's behavior or some other sort of event of condition. This is the kind of question that pilot testing indicated was feasible.

The next step was to decide what population of reinforcers to use for the questions. One can, of course, generate a population of reinforcers without any prior explicit definition of the domain he wishes to sample. However, it is preferable where possible to define the class of events to be sampled--here, to represent as reinforcers. In this the literature regarding behavior modification, especially in classroom use (e.g., Becker, Thomas, & Carnine, 1969), as well as Rotter's (1954) point of view, were followed. It was assumed that, for preschool age children, the primary reinforcers--at least, those of most importance in school and in socialization--are the attention and approval of parents and teachers, and perhaps to a lesser extent of peers, and perhaps also some sorts of self-approval. More specifically, it was assumed that it is observable cues of approval and attention,

or events associated with increased probability of approval and attention, that serve as reinforcing events. These cues of approval and attention, then, could constitute the questions for a free-response measure of IE; and they could be expressed simply enough to stay within preschoolers' language limits. So the questions that were tried, which pilot testing indicated were effective, were questions like "What makes mothers smile?" and "What makes you happy?" Preschoolers had little problem understanding such questions. Children did often cite some behavior of their own in response to these questions, reflecting an association and therefore presumably expectancy of relation between that reinforcement and that behavior; and children also often cited something other than their behavior, like the weather, or other people's behavior, and so forth. For example, to "What makes mothers smile?" there were answers like "When I draw her a picture" and "When I set the table," clearly reflecting perceived internal control contingencies, and also answers like "When daddy comes home" and "When she buys new shoes," clearly reflecting "external" contingencies.

Once informal pilot testing had established that such questions seemed feasible with disadvantaged preschool-age children, an initial study was designed to evaluate the rater reliability for coding responses as to whether they reflected internal or external control expectancies. The same study was designed to provide also an indication of retest reliability of scores devised by this method and some evidence of convergent validity of the method, by (1) generating two separate sets of questions, to be given at separate times, as alternate forms of this measure, and (2) attempting to adapt some questions from previous IE tests for use as an external test criterion. A final purpose of this initial study was to investigate whether responses were biased, among disadvantaged black children, by the ethnicity of the interviewer.

METHOD

Test Items.

A set of 40 questions was generated, fitting a 2 X 5 (Reinforcement Type X Reinforcement Agent) model with four questions per category. These questions collectively are called the Stephens-Delys Reinforcement Contingency Interview (SDRCI). Half the questions concerned negative reinforcement, half positive reinforcement (the reinforcement type factor). This permitted subscaling for positive vs. negative IE expectancies, as do the IAR and Gruen tests, and obviated the confounding of IE with expectancy of positive reinforcement (Crandall, 1971). The five reinforcement agents were self, peers ("other kids"), mothers, fathers, and teachers. ("Mothers" and "fathers" were used rather than "your mother" and "your father" for two reasons: to avoid problems occasioned by fatherless or motherless children, and to slightly "depersonalize" questions so as to minimize objections regarding invasion of privacy.) Reinforcement agent subscales were, thus, also available, either across or within reinforcement type and for either single reinforcement agents or combinations (e.g., both parents or all three adults). The specific questions used are given in Table 1.

Table 1

SDRCI Questions

1. What makes you happy?	21. What makes you smile?
2. What makes other kids happy?	22. What makes other kids smile?
3. What makes mothers happy?	23. What makes mothers smile?
4. What makes fathers happy?	24. What makes fathers smile?
5. What makes teachers happy?	25. What makes teachers smile?
6. What makes you unhappy?	26. What makes you look mean?
7. What makes other kids unhappy?	27. What makes other kids look mean?
8. What makes mothers unhappy?	28. What makes mothers look mean?
9. What makes fathers unhappy?	29. What makes fathers look mean?
10. What makes teachers unhappy?	30. What makes teachers look mean?
11. What makes you nice?	31. What makes you feel good?
12. What makes other kids nice?	32. What makes other kids like you?
13. What makes mothers be nice?	33. What makes mothers love you?
14. What makes fathers be nice?	34. What makes fathers love you?
15. What makes teachers be nice?	35. What makes teachers like you?
16. What makes you be not nice?	36. What makes you angry?
17. What makes other kids not nice?	37. What makes other kids angry?
18. What makes mothers be not nice?	38. What makes mothers angry?
19. What makes fathers be not nice?	39. What makes fathers angry?
20. What makes teachers be not nice?	40. What makes teachers angry?

Testing Procedure.

To communicate the intent of the questions, the specific questions sometimes had to be reworded (e.g., "mad" for "unhappy") and repeated, especially with children of minimal language development, and supplemented with facial expression and inflection cues. The procedure, then, did not entail adhering rigidly to the narrow semantic content of each specific question, but asking, as it were, "What makes positively (or negatively) reinforcing behavior by ___ more likely to happen?" four times in four different ways. (It was not uncommon for children to complain, "But you already asked that," on later questions, even when the specific wording of questions had been maintained.) Similarly, it was not required that the questions be given in the prescribed order, although departure from their order seldom seemed to help much in interviewing a difficult child and pilot testing indicated that the question order cited seemed easier for most children than alternative orders.

Subjects.

Subjects were 14 female and 10 male black Head Start children who had completed one year of a four-year-old Head Start program, were attending a summer session, and were age-eligible for kindergarten the following fall.

Measures.

The interview questions were divided into two alternate forms, Form A and Form B (items 1-20 and 21-40 as listed). After the child responded to each question, he was asked to give a second response (e.g., "That's fine. Now, what else makes you happy?"). The purpose of this was to increase the number

of responses on each form to 40; the purpose of constructing alternate forms was to provide some evidence of a parallel-forms type of convergent validity.

In an attempt to provide further convergent validity evidence, 30 questions were adapted from the IAR, the NSIE, and the Bialer-Cromwell IE tests. These were the questions that seemed most likely to be meaningful and comprehensible to preschoolers, reworded to simplify them as much as possible. Fifteen of the questions were yes-no format, adapted from the NSIE and Bialer-Cromwell tests, and 15 were forced choice, adapted from the IAR. They, too, were administered in an individual interview fashion, at a separate session and by an interviewer other than the SDRCI interviewers.

Interviewers.

For the SDRCI interviewing, one interviewer was a 16-year-old black girl, the other interviewer a young white woman. Approximately half the subjects were given Form A by each of these interviewers. Form B was given about two weeks later; for Form B, approximately half the children switched interviewers.

Date Analysis.

All 48 SDRCI protocols were scored independently by two raters, without use of a detailed item-by-item scoring manual per se but with a set of general scoring rules. Correlations were computed to determine rater reliability. Scores on Form A and Form B were correlated as a lower bound index of two-week retest reliability and simultaneously as an index of convergent validity as reflected by correlation between parallel forms. SDRCI scores (separately for each form and also combined scores) were correlated with scores on the questionnaire-type items (all items combined and also IAR-type and yes-no type items separately). Correlations were also computed for positive reinforcement vs. negative reinforcement items, and also among the five subscales defined by reinforcement agent, both within and across forms.

An unweighted means 2 X 2 X 2 X 2 X 2 X 5 factorial analysis of variance was performed, primarily to investigate interviewer effects on frequency of internal control responses. The dependent variable was number of internal control responses. Independent variables were interviewer on Form A (Interviewer I, black female adolescent; Interviewer II, white young woman), interviewer on Form B, and (as repeated measures factors) form (A vs. B), reinforcement type (positive or negative), response number (first vs. second response on each question), and reinforcement agent.

RESULTS

Rater reliability for total scores was .98. The correlation between scores on the two forms was .69. On the questionnaire-type measure, scores on the IAR-type items actually correlated negatively ($r = -.19$) with scores on the yes-no type items; during testing it was obvious that both types of items were, despite all efforts, still too difficult for the children to respond to meaningfully. SDRCI scores, combining Forms A and B, correlated .29 (.23 for Form A alone and .32 for Form B alone) with combined scores on the questionnaire measure, .17 with the IAR items alone, and .29 with the yes-no items alone; none of the

correlations involving questionnaire items approached significance.

Among SDRCI subscales, positive reinforcement items correlated with negative reinforcement items .83 on Form A and .79 on Form B; on both forms there were higher correlations among mother, father, and teacher items (all correlated .75 to .83 with total SDRCI scores) than self and peer items (which correlated .21 to .60 with total scores).

Interviewer differences had no significant effect on overall scores. One interviewer tended to accept more "Don't know" responses (which are scored external), especially on the second response where those were more common anyway, than did the other interviewer. This produced slightly but nonsignificantly lower internal control scores for the children she interviewed.

There were higher scores on mother, father, and teacher items than on self and peer items (reinforcement agent main effect, $F = 17.1$, df 4/80, $p < .001$). Internal control scores were significantly higher on the first response than the second ($F = 14.5$, df 1/20, $p < .001$), almost totally a result of the greater frequency of "Don't know" responses on second than first responses.

DISCUSSION

Rater reliability was, obviously, adequate. There were some responses on each protocol that could not confidently be coded either internal or external, but these did not affect total scores enough to jeopardize reliability; and most responses were obvious internal or external responses. A scoring manual was prepared subsequently to establish conventions for scoring the ambiguous responses and also to provide an expanded operational definition of IE as defined in this method (see Appendix 1).

Evidence concerning convergent validity was less clear. The correlation of .69 between parallel forms administered two weeks apart reflected convergent validity across items (and time) but within method. The two types of conventional test items used did not correlate significantly with SDRCI scores; but, given the evidence of lack of meaningfulness of these items in this population, no clear conclusion can be drawn from these data. (Two subsequent studies, one with second graders (Stephens, 1971a, 1972b) and one with 83 third graders (Stephens, 1972a), have intercorrelated SDRCI, NSIE, Gruen test, and, in the third graders' study, IAR scores. These studies have found (1) generally quite low correlations among all four of these tests at these ages, (2) sex differences in regard to the pattern of intercorrelation among tests, and (3) the Gruen test showing generally the highest intercorrelations with other tests (although still seldom higher than the .30s, and often lower) and the SDRCI next highest. The most obvious and important implication of these data is that these four tests reflect different variables, even though each variable might justifiably be called "locus of control." Inspection of the items, format, and reported correlates of these tests, as discussed elsewhere (Stephens, 1971a, 1972b, 1972a), supports the conclusion that the tests reflect different variables, albeit with many shared properties and correlates, which refer to different connotative meanings the term "locus of control" has assumed in the literature.)

(Several studies have subsequently yielded further information regarding reliability and validity of the SDRCI. In all these studies, Forms A and B have been combined, and only one response per question has been elicited. One study (Stephens, 1972a) found four-month retest reliability of .62 in a sample of 49 white middle-class nursery school children, to supplement the .69 two-week retest reliability coefficient in this initial study. Another study (Stephens, 1971a, 1972b) found a Kuder-Richardson 20 coefficient of internal consistency of .82 in a sample of 575 second graders. The high correlations among subscales further support the reliability of the test and indicate more generality of individual differences in IE than might have been expected (see Mischel, 1968); they also, however, indicate that the subscales may be of limited discriminative utility. On the other hand, factor analyses and item analyses (Stephens, 1971a; Muller, 1972) of the second graders' data indicated that self-reinforcement items reflect a different variable than that represented by the other four reinforcement agents; and in one study (Stephens, 1971a, 1972b), self-reinforcement and teacher-reinforcement items were differentially sensitive to differences in early elementary educational experiences.)

Subsequent studies have also yielded more data regarding interviewer effects. One study (Stephens, Delys, Lopez-Roig, & Vilez, 1971), with Puerto Rican and Chicano children, did find an interviewer ethnicity effect on IE response content among these children; but in no other study have interviewer effects been apparent. In one study (Stephens, 1971a), approximately half of the 575 second graders tested were given the first half of the questions by one interviewer and the second half by another; these children showed the same correlation (.61) between first-half and second-half scores as did the children given all questions by the same interviewer, and between-interviewer differences in mean scores were nonsignificant (Stephens, 1971a). It would appear possible that interviewers might, unintentionally, selectively reinforce internal or external responses. Interviewers are warned of this possibility when trained and monitored to control this possible artifact; there has been no evidence that such a bias has yet occurred.

It has been found that the major problem in interviewing is the tendency for children to repeat the same response to consecutive questions. This perseverance is, of course, most common among relatively young, nonverbal, frightened, and/or bored children. The procedure that was adopted to combat this is to accept a repeated response only once, then to persist with a comment such as, "Yes, that's fine; now let's think of something else that makes teachers happy too." Nevertheless, young and difficult-to-interview children tend to repeat, if not the same response verbatim, then responses of the same class. Following is an illustrative interchange: "What makes mothers happy?" "Kitties." "What makes fathers happy?" "Kitties." "That's fine. Now let's think of something else that makes fathers happy." "Puppy dogs." The task of thinking of a response is difficult enough that, for some children, having thought of a response of one class (either an internal-type or an external-type response) seems to be self-reinforcing and to serve to increase the probability (and decrease the latency) of responses of that class subsequently. This phenomenon seems to spuriously increase the range of SDRCI scores and decrease its sensitivity as a measure of individual differences among young children. Another occasional problem is "nonsense" responding (e.g., "What makes teachers happy?" "Chairs."), again from difficult-to-interview children. Recent or impending salient events, such as Christmas or birth of a sibling, also obviously affect responses, as would

be expected in an associative-type technique; but they do not necessarily bias the internal vs. external control aspect of the responses. The measure could probably be made more powerful if the response-repetition and nonsense response tendencies could be eliminated. Nevertheless, the measure seems to have enough power to assess differences between groups and, with sufficient sample size, to assess individual differences correlates within groups as well.

The most common avoidable problem in scoring is a response in which the child cites an ambiguous "someone," who may and may not be the child himself, as the actor: e.g., "What makes teachers angry?" "When someone won't keep still." Sufficient probing can usually clarify whether the child is referring thus to himself, to another child, or to children in general (in which case the response is scored internal, since the subject is himself a member of the group--children--denoted). Subsequent experience has also indicated that "Don't know" responses can be essentially totally avoided if probing is sufficiently persistent.

Despite the problems in interviewing and scoring, this initial study (and subsequent ones) indicated that the SDRCI is sufficiently reliable and unbiased by interviewer ethnicity to be used to investigate IE in disadvantaged preschool children. Evidence of convergent validity was obscured because of the failure of the yes-no and forced-choice test items to elicit reliable data.

STUDY II: INITIAL CONSTRUCT VALIDATION

Introduction

The second study was designed to provide construct validity evidence for the SDRCI. To assess construct validity, a number of IE correlates (other than scores on other IE tests) might be used. They could include either antecedents (e.g., socioeconomic status, age, or parent attitude antecedents implicated by previous research) or concurrent or subsequent variables (e.g., performance on various tasks). Rotter (1966), Joe (1971), and others have reviewed the literature regarding such correlates of internal control expectancies as, for example, personality test scores, performance in social influence or conformity tasks, etc. However, the most direct and relevant IE correlate, for both theoretical and practical reasons, is, simply, response to reinforcement: subjects with more internal control expectancies ("internals") should show more modification of their behavior in response to reinforcement than do subjects with more external control expectancies ("externals"). Specifically, internals' behavior should show more distinct and rapid shaping to reinforcement contingencies than externals' behavior. This relationship is the primary definition of IE in Rotter's (1966) system. It is also the property of IE with greatest educational relevance: its contribution to learning and performance. Consequently, this study sought to test the relationship of IE differences, as assessed by the SDRCI, to learning, in a controlled task with clear reinforcement contingencies.

In designing the study it was assumed that the task ought optimally to be nonverbal, and also to require only minimal ability to follow verbal instructions, so that verbal ability differences would contribute minimally to performance differences. It was also assumed that the task ought optimally not to require authority-mediated reinforcement, like verbal approval - again,

so that subject differences in responsiveness to adult approval per se would be of minimal influence. Similarly, it was desirable not to have to rely on material reinforcers such as candy. The task, then, needed to provide direct evidence to the child himself as to whether his response was correct (or "good") or not, and needed to be of enough intrinsic interest that evidence of good performance itself might be an adequate reinforcer for the child. The task ultimately selected to meet these criteria was a mirror tracing task, providing the child a number of trials in which he could learn how better to use the mirror-reversed visual feedback to enable him to trace a pattern quickly without his pencil going outside the alleyway.

In investigating the role of IF in performance of this task, however, another variable had to be considered, for both theoretical and practical reasons: aggressiveness. Hudgens and MacNeil (1970) had found that mice with a history of success in aggressive encounters perform less well on a stressful avoidance learning task than do mice with a history of failures. Similar findings with monkeys had been reported by Levine and Gordon (1968). These authors suggested that, to the extent that a subject has been reinforced for (successful in) aggressive responses in stressful situations in the past, he will tend to persist with aggressive responses in new situations, so that his acquisition of other responses will be disrupted. Whether a similar phenomenon occurs in humans, and particularly children, had not been tested. However, in many compensatory preschool programs (and even nursery schools for middle-class children) there are children of markedly aggressive disposition; these children often, according to teacher reports, appear to profit little from the cognitive training in the preschool program; and, for this and other more obvious reasons, aggression training absorbs much of the effort of the teachers. Furthermore, Crandall (1971) found IE related to aggressiveness in childhood. If aggressive disposition does interfere with performance on learning tasks, it could obscure the relation of IE to performance.

The present study was designed, therefore, to test the relation of both IE and of aggressive disposition to performance on an instrumental learning task.

METHOD

Subjects.

Thirty-two preschool age (39-68 months) boys in a day care center were subjects; all were of essentially lower-middle-class or middle-class socio-economic status.

Dependent Variable Task.

Pilot testing revealed that few children of preschool age could perform mirror-tracing designs with angles, nor even with diagonal or horizontal lines; the only such task within the abilities of most preschoolers proved to be a single vertical pathway. The pathway adopted was 5-1/2 in. long by 1/4 in. wide. Each subject had 10 trials, each on a separate score sheet (see Parker [1971] for further details of apparatus and procedure); a small felt-tip pen was used. Errors were defined as lifting the pen, going backwards in the pathway, or going outside the pathway; when the subject went outside the pathway, his pen was put back into the pathway at the same point it went out. Number of errors and elapsed time in seconds were recorded on each trial as the performance measure

Independent Variables.

The subject's SDRCI score, as the IE measure, was the independent variable of primary interest; all 40 SDRCI questions were used, eliciting just one response per question. In addition there were three measures of individual differences in aggressiveness and an experimental manipulation of state-type aggressive dispositions. Subjects were assigned randomly (within each IE score group) to an Aggression-Modeling (AM) or Nonaggression-Modeling (NM) condition. In the AM condition the subject was brought to the experimental room and confronted with an array of toys typically associated with aggression (a large inflatable "bop bag" and cowboy guns and hat) and nonaggressive toys (Lego building blocks and colorful plastic discs). He was told by the experimenter that these were some of the toys children "at the school I go to" play with, and was then shown a 45-sec videotape of a boy playing aggressively with the toys: noisily kicking and punching the bop bag, shooting the guns, etc. Toward the end of the videotape the experimenter was shown entering, approvingly, and telling the model it was time to go. The subject was asked to perform the learning task immediately after viewing the videotape; he was allowed to play with the toys thereafter. In the NM condition the procedure was identical except that the model was shown playing quietly and gently with the nonaggressive toys.

To provide the primary measure of trait-type individual differences in aggressiveness, a 7-point teacher-rating scale of aggressiveness was developed (see Parker, 1971), in consultation with the teachers. It employed extensive behavioral definitions of the extremely aggressive (e.g., "... is rough, ... hitting, ... pushing, shouting, kicking, ...") and extremely nonaggressive (e.g., "... will not fight back, ... ignores the aggressive peer, ... cry, ...") ends of the scale and of the midrange of the scale. The head teacher rated all 32 subjects, and each subject was also rated by his own teacher, to permit an estimate of rater reliability. To provide an external criterion of the validity of these teacher ratings, a 10-item peer nomination measure was also devised, with five aggressive items (e.g., "Which boy in your class fights the most?") and five nonaggressive items (e.g., "Which boy in your class doesn't fight at all?"); each subject named one child for each question.

The third index of aggressive response tendency was drawn from SDRCI protocols. SDRCI responses can also be classified not only as to whether they reflect internal or external control expectancies, but also as to whether they reflect overt aggression. Number of aggressive responses on the SDRCI was recorded as the third measure of aggressive disposition. The SDRCI was used, then, primarily as the IE measure, but also to provide this index of aggressive response tendencies.

Procedure.

SDRCI testing was completed first, over a period of several weeks. SDRCI protocols were scored to define above- and below-median IE groups. Then mirror-tracing testing was done, over a period of several days, by the same experimenter with subjects within each IE score group assigned randomly to aggression-modeling condition. The subject first viewed the video-tape (AM or NM), then performed the mirror-tracing task, then was allowed to play with the toys; after this he was asked the peer-nomination aggression-rating questions. Teacher ratings of aggression and scoring of the SDRCI for aggression responses were done after all the subjects had been tested; both the teacher ratings and the peer nomination measure had been developed through prior pilot testing.

Data Analysis.

The primary analyses (one with number of errors per trial and one with elapsed time in seconds per trial as dependent variable) were 2 X 2 X 10 split-plot design univariate analyses of variance (Winer, 1962). Independent variables were IE score (above vs. below the median internal control score), aggression-modeling condition (AM vs. NM), and, as a repeated measures factor, trials. Because of the skewing of scores (toward zero error and low time scores), a median test (Siegal, 1956) was also performed. Analyses were repeated with subjects classified as to teacher ratings of aggressiveness rather than IE score.

In addition, rater reliability of teacher ratings was assessed, as was rater reliability of scoring SDRCI responses aggressive vs. nonaggressive; and the three aggressiveness measures were intercorrelated with one another and with IE scores.

RESULTS

Individual differences in IE were significantly related to performance. "Internals" (subjects with above-median SDRCI internal control scores) had both lower time scores ($F = 6.4$, $df = 1/28$, $p < .05$) and fewer errors ($F = 4.2$, $df = 1/28$, $p < .05$) than "externals." There were no significant interactions, on either the time or the error criterion, with modeling conditions or trials. IE scores were also significantly related to both teacher ratings and peer ratings of aggression, as well as number of aggression responses on the SDRCI ($p < .05$ in each); in each instance, higher aggressiveness scores were associated with higher internal control scores.

The aggression-modeling manipulation, however, had no effect on errors ($F < 1$), nor was the effect on time scores significant ($F = 2.3$, $df = 1/28$, $p > .05$). (The nonsignificant trend was, however, as predicted: time scores were higher for AM than NM children.) When they were allowed to play with the toys after performing the tracing task, the subjects often did manifest the specific aggressive (or nonaggressive) behaviors they had seen modeled; but the modeling did not seem to produce any clear heightening of arousal or excitement, nor did it affect performance significantly. There was also no relation of performance to teacher ratings of individual differences in aggression.

The median tests produced similar results: "internals" performed significantly better than "externals" on both time ($p < .001$) and error ($p < .05$) measures, with no significant differences attributable to teacher ratings of aggressiveness or to aggression modeling. The correlation (Pearson r) between SDRCI scores and error scores was $-.34$ ($p = .06$), and between SDRCI scores and time scores $-.35$ ($p < .05$).

Rater reliability of teacher ratings was .81, of scoring SDRCI responses for aggression .99. It was apparent, however, that peer ratings were contaminated by a salience-conspicuousness factor: when a subject named a particular child for an aggressive item (e.g., "Which boy are you most scared of"), he often named the same child on the opposite item ("Which boy are you never scared of"). Teacher ratings correlated .63 with peer nomination scores based on the aggressive items alone, but nonsignificantly with scores based on aggressive and nonaggressive items combined. SDRCI aggression scores correlated

only .30 ($p > .05$) with teacher ratings, and had the same correlation with peer ratings on aggressive items: SDRCI responses did not appear, that is, to reflect overt aggressive dispositions.

DISCUSSION

SDRCI scores were related to performance as predicted, thus providing evidence of construct validity of the SDRCI as an IE measure. However, the relation could have been attributable to differences in age or intelligence. Chronological age correlated .51 ($p < .01$) with SDRCI scores in this sample, -.19 ($p > .20$) with time, and -.31 ($p = .08$) with error scores. Controlling for age differences, the partial correlation of IE with error scores was -.25 ($p = .15$), and with time scores -.38 ($p = .05$). For 20 of the 32 subjects there were available scores on the Slosson Intelligence Test (SIT). For these 20 children SIT IQ scores correlated .39 ($p = .09$) with IE, -.58 ($p < .01$) with error, and -.59 ($p < .01$) with time scores. Controlling for these IQ score differences, the partial correlation of IE with error scores was only .01 and with time scores -.14 ($.30 > p > .20$) for these 20 children.

(A subsequent study (Stephens, 1972a) was designed in part to replicate the present study and in part to further investigate the role of age and intelligence in mediating the relation of SDRCI scores to performance on this task. There was no manipulation of aggression-modeling conditions or measure of individual differences in aggression, but a larger sample of children with intelligence scores available and a more restricted age range. Each subject was given the SDRCI, the mirror-tracing task, and--among other tests and tasks--the Peabody Picture Vocabulary Test (PPVT), the Raven Progressive Matrices (RPM), and the Picture Completion (PCom) subtest of the Wechsler Preschool and Primary Scale of Intelligence. Subjects were the 50 children (both boys and girls) in two parent cooperative nursery school classes and one class from a university laboratory nursery school. SDRCI scores correlated -.34 ($p < .05$) with error and -.14 ($p > .30$) with time scores, thus replicating the first study significantly with the error and nonsignificantly but in the same direction with the time criterion and finding that the relation of SDRCI scores to mirror-tracing performance is not sex specific. SDRCI scores correlated -.23 ($p > .05$) with PPVT, -.05 with RPM, and .01 with PCom scores; error scores correlated only -.04, .05, and -.05, respectively, with these tests, and time scores .02, .01, and .12. Age differences correlated only .06 with SDRCI scores, .24 ($p > .10$) with error, and .28 ($p > .05$) with time scores. In this sample, then, age and intelligence differences could not account for the relation of SDRCI scores to performance on the mirror-tracing task; so the relation of SDRCI scores to instrumental learning independent of correlated age or intelligence differences was supported.)

On balance, then, the construct validity of the SDRCI as an IE measure was supported, and direct evidence found of a relation of IE to performance on a learning task among preschool children. The suggestion in the initial construct validation study that the relation of SDRCI scores to performance may have been an artifact of age and intelligence correlates of SDRCI scores was not supported in the subsequent replication, but the subsequent study did confirm the relation of SDRCI scores to error scores on this task (although the relation of SDRCI to time scores in the latter study fell short of significance).

STUDY III: SOCIOECONOMIC DIFFERENCES AMONG PRESCHOOL AGE CHILDREN

Introduction

The next study was designed to test whether socioeconomic differences in IE appear before school age. Despite the evidence of socioeconomic differences in IE among elementary school and older children, it is entirely possible that the relatively external expectancies of disadvantaged children do not develop before second or third grade level; they might be, in fact, the result, rather than an antecedent, of unsuccessful school experience. This, then, is a crucial question for determining the optimal emphasis and curriculum for compensatory education programs for preschool children. The present study sought to investigate both ethnic and economic status differences in IE among preschoolers, comparing one black and one white economically disadvantaged group with each other and with two separate groups of white nondisadvantaged children.

METHOD

Subjects.

Disadvantaged subjects were children in two black and two white Head Start classes; there were 12 girls and 8 boys in the white group and 17 girls and 18 boys in the black group. Nondisadvantaged subjects came from two parent cooperative nursery schools (19 girls and 15 boys, all white) and a private Montessori preschool (7 girls and 9 boys, all but one white). All came from the same city, and at the beginning of the school year they were within one year of the age required for beginning kindergarten.

Measure.

The SDRCI measure of locus of control expectancies was the only measure taken; again, all 40 questions were used, one response per question.

Procedure and Data Analysis.

All children were tested within a month after the beginning of the school year. Testing was done away from other children or adults, in hallways or extra rooms, as available, at the site of the child's classroom.

To test differences between advantaged and disadvantaged children, an unweighted means factorial 2 X 2 X 2 X 5 analysis of variance of internal control responses was performed. Independent variables were economic status, sex, and, as repeated measures factors, reinforcement type (positive vs. negative) and reinforcement agent (mothers, fathers, teachers, etc.). Subsequent a priori 2 X 2 X 2 X 5 ANOVs were performed to test differences between black and white Head Start children and between Montessori and parent cooperative middle-class nursery school children.

RESULTS

Middle-class subjects had significantly higher total internal scores than did disadvantaged subjects ($F = 8.24$, $df = 1/83$, $p < .01$). Means for boys and girls in each of the four separate groups are in Table 2. Differences between

black and white Head Start groups were nonsignificant ($F < 1$), as were differences between Montessori and parent cooperative nursery school middle-class groups.

Table 2

Mean SDRCI Internal Control Scores for Disadvantaged and Middle-Class Groups

	Disadvantaged		Middle Class	
	Black	White	Montessori	Parent Cooperati
Girls	8.7	13.0	15.8	14.9
Boys	8.1	8.5	14.4	11.5

Girls tended to have higher internal control scores than boys, but the difference was nonsignificant ($.20 > p > .10$). However, the interaction of sex with reinforcement type was significant ($F = 6.11$, $df = 1/82$, $p < .02$): girls' scores on positive reinforcement items were higher than their scores on negative reinforcement items or boys' scores on either positive or negative reinforcement items. The reinforcement type main effect was significant ($F = 9.66$, $df = 1/82$, $p < .01$), scores being higher on positive reinforcement than on negative reinforcement items, as was the reinforcement agent main effect ($F = 16.48$, $df = 4/328$, $p < .001$): scores were highest overall on teacher items and lowest on peer items, but the patterning of differences between reinforcement agent subsca varied considerably (if nonsignificantly) with economic status, sex, and reinforcement type.

DISCUSSION

The results confirmed that disadvantaged children do manifest relatively external control expectancies already by preschool age. The first study, using a slightly different SDRCI procedure with 24 black Head Start children at the end of the school year, found scores in the same range as the means for the Head Start groups in this study. A subsequent study (Stephens, Delys, & Parker, 1971) found that even among children in a non-Head Start compensatory program, those below the OEO "poverty" line had lower ($p = .08$) internal control scores than those above-- if only marginally above--the "poverty" line; again those below the "poverty" line had scores in the same range as those reported above, whereas means for the subjects above the "poverty" line were intermediate between those scores and the means for middle-class groups reported here. Other studies (see Stephens, 1971b) have found scores for other white and black disadvantaged groups, and for white middle-class groups, consistent with these findings. These differer in IE expectancies, then, do precede school experience; and it would indeed appear that compensatory preschool programs might profitably be aimed specifically at enhancing development of internal control expectancies.

STUDY IV: INITIAL ASSESSMENT OF IMPACT ON IE OF DIFFERENT TYPES OF PRESCHOOL PROGRAMS

Introduction

The final study was designed to provide an initial comparison of the effects of different kinds of preschool programs on development of internal control expectancies. This is, of course, the question of ultimately the greatest interest: how can preschool experiences best be designed to enhance IE development? This study could only represent an initial inquiry into the question, however. The size of the project made it impossible to exert proper control over such variables as socioeconomic and cultural differences, parent attitude and cooperation differences, teacher training and effectiveness differences, teacher-pupil ratios, etc., to permit really adequate tests of the effects of curricular and program differences. It was not politically feasible to assign subjects randomly to programs; sample sizes of children were relatively small, especially for comparisons involving change scores; and the number of teachers, or classrooms, in each program sampled was even much less, making even more tenuous generalizations regarding such programs. Nonetheless, this study was designed to provide at least an initial comparison of several kinds of preschool programs.

Of particular interest were two programs: a more or less "typical" Head Start program and a Montessori program, the latter chosen because of its emphasis on self-directed learning, etc. - i.e., on what would appear to be "internal control expectancy training" regarding learning and school performance. One of the other programs was a parent cooperative nursery school, chosen because its emphasis was not so much on cognitive development, either in the Montessori sense or as in compensatory programs, but rather on parent education, with apparently relatively greater emphasis on social and emotional development and less on cognitive development than in either Montessori or compensatory programs. The fourth program investigated was a non-Head Start but federally funded compensatory program. It differed from the Head Start program in several ways: it was much more generously funded, had more highly trained teachers (all held at least a bachelor's degree and a teacher's license and had prior teaching experience), had a more favorable teacher-pupil ratio (one teacher and two aides per class of about twelve children), etc.; perhaps more relevant, its program was planned in much more detail, and involved substantially more time in organized, structured, direct-teaching activities (3-5 children per teacher), than did any of the other programs.

The purpose of this study was, then, to compare these four programs with one another in regard to the amount of increase in internal control expectancies, from the beginning to the end of the school year, shown by the children in each program.

METHOD

Subjects.

Subjects were the children enrolled in these four programs. There were five boys and 15 girls (all black and all disadvantaged) in the Head Start group, six boys and six girls (one black, the rest white), in the Montessori group, nine boys and 13 girls (all white, all middle class) in the middle class parent cooperative nursery school ("Co-op") group, and 27 boys and 33 girls (all black, all below or near the OEO criterion for "disadvantaged") in the non-Head Start

("Compensatory") group. The total sample numbered 114. This reflected substantial attrition from the sample tested originally, attributable to children's having left their respective schools during the year and/or to absence at time of testing. The groups represented one Montessori class, two Head Start classes, two Co-op classes, and eight Compensatory classes.

Measure.

The SDRCI was the only measure given to all subjects, although other tests were available for the Compensatory subjects; again, all 40 questions were given, one response per question.

Procedure and Data Analysis.

All subjects were tested at the beginning of the school year (between late September and late October) and again at the end of the school year (in May or early June). In all groups pretesting and post-testing were done in empty rooms and/or hallways; testing conditions were adequate.

Data were analyzed by means of a 4 X 2 X 2 X 2 X 5 repeated measures analysis of variance (unweighted means). The dependent variable was number of internal control responses. There were two between-groups independent variables: program (Head Start, Montessori, Co-op, and Compensatory) and sex. The other three independent variables were testing time (pre or post), reinforcement type (positive vs. negative reinforcement questions on the SDRCI) and reinforcement agent. The data of primary interest were those relevant to the interaction between program and testing time, which would reflect whether programs differed significantly in regard to amount of change in SDRCI scores from beginning to end of year.

RESULTS

The program X testing time interaction was nonsignificant ($F = 1.02$, $df = 3/106$, $p = .39$). The trend of differences was, however, generally as expected: the Montessori group showed most increase (from 16.3 to 23.0); the Compensatory group had slightly less increase (from 11.5 to 13.9) than did the Head Start group (from 7.7 to 10.3); and Co-op children had a mean increase intermediate between that of the Montessori and of the two disadvantaged groups (from 12.6 to 16.5).

There was a significant 4-way interaction, however, involving program and testing time and also sex and reinforcement agent ($F = 2.2$, $df = 12/424$, $p = .01$). There was no simple pattern apparent which could explain or describe this interaction, however (see Table 3).

The data of primary theoretical interest are the pre- to post-test differences for each group for two particular reinforcement agents: teacher and self. The Montessori boys actually showed a slight decrease in scores on self items, but a large increase on teacher items, whereas the Montessori girls showed only a slight increase on the teacher items but a large increase on self items. Other groups showed other sex differences, but no general pattern or explanation was apparent. No subsequent simple effects analyses were conducted because the absence of any meaningful pattern, combined with very small sample sizes and numerous uncontrolled sampling variables, made such analyses specious.

Table 3

Pretest-Post-test Scores for Each Sex in Each Program

on Each SDRCI Reinforcement Agent Subscale

Subscale	Program							
	Montessori		Head Start		Compensatory		Co-op	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Self Pretest	3.2	2.2	1.6	2.3	2.6	3.2	3.7	3.0
Self Post-test	3.0	4.5	3.2	2.8	2.9	2.8	3.8	4.1
Peer Pretest	1.3	2.2	0.6	1.5	1.2	1.8	1.3	2.0
Peer Post-test	5.5	3.0	1.2	1.0	1.9	2.4	1.8	3.1
Mother Pre-test	3.7	3.7	1.6	1.2	1.7	2.1	1.8	2.5
Mother Post-test	6.3	4.5	1.8	2.4	2.8	3.0	2.8	3.6
Father Pre-test	4.1	3.8	1.0	0.7	1.4	2.1	1.3	2.9
Father Post-test	6.5	4.0	0.8	1.7	2.2	2.2	1.9	3.5
Teacher Pre-test	4.5	4.0	1.6	3.3	3.4	3.4	2.9	3.8
Teacher Post-test	6.2	4.3	3.0	2.7	3.4	4.0	3.4	5.0

The analysis did reveal significant differences among groups ($F = 12.1$, $df = 3/106$, $p < .001$), and a significant overall increase from pre- to post-test ($F = 12.0$, $df = 1/106$, $p = .001$); it also revealed significantly higher scores on positive than on negative reinforcement items ($F = 3.7$, $df = 1/106$, $p = .05$), significant differences attributable to reinforcement agents ($F = 16.9$, $df = 4/424$, $p < .001$), and significant interactions of the reinforcement agent factor with both program ($F = 2.8$, $df = 12/424$, $p = .001$) and reinforcement type ($F = 2.9$, $df = 4/424$, $p = .02$). None of these was of theoretical interest, however.

DISCUSSION

No confident conclusions can be drawn from these findings, and all discussion must be very careful, because of the small sample sizes, attrition, uncontrolled sampling variables, etc., as well as because of the nonsignificance of the differences found. It would be as incautious to conclude, because the differences were not significant ($p = .39$), that the programs do not differ in their impact on IE as it would to conclude that they do. The nonsignificant differences which were found are in fact consistent with a priori expectation: the Montessori program produced the greatest increase in SDRCI scores; the Compensatory program actually produced slightly less increase than did the Head Start program despite its much more generous support, much better trained and equipped teachers, etc.; and the Co-op program produced slightly greater increase than did either Compensatory program. Clearly there is need for a subsequent study to test more definitively whether these and/or other preschool programs differ in their impact on IE. However, such a study must have both the political and the financial resources to permit such control procedures as random assignment to program and permit sampling a number of teachers (or classrooms) per program.

(A subsequent study [Stephens, 1971a, 1972b] has investigated the impact on IE of two different Project Follow Through programs: the EDC-sponsored "open classroom" model and the Engelmann-Becker "behavior modification" type of program. A total of 575 children, in 27 classrooms, were tested, including two different EDC and two different Engelmann-Becker programs (one each with predominantly white and one each with predominantly black children). In each group the children had been in their respective programs for kindergarten, first, and second grade. Both Engelmann-Becker programs produced higher SDRCI scores than found in a non-Follow Through comparison group of black disadvantaged children, but EDC programs produced higher scores than Engelmann--Becker programs; a middle-class white "open classroom" (non-Follow Through) group showed highest scores of all. Differences between groups were statistically significant, although in some instances not large. It appears, then, that although cultural or economic status effects may persist despite school experiences, less structured early educational programs which emphasize child-directed learning experiences may enhance internal control development more than more structured ones, with more teacher-directed learning. However, there were substantial differences among classes within programs, as well as among children within classes; and there were also substantial differences among teachers within programs in teacher-pupil interaction measures. It is still unclear, then, what are the specific aspects of teacher behavior and other characteristics of the program which may be instrumental in enhancing internal control expectancy development and thereby effective in compensatory programs aimed at either preschool or early elementary level.)

Bibliography

Atkinson, J. W. An Introduction to Motivation. Princeton, New Jersey: Van Nostrand, 1964.

Battle, Esther S., & Rotter, J. B. Children's feelings of personal control as related to social class and ethnic groups. Journal of Personality, 1963, 31, 482-490.

Becker, W. C., Thomas, D. R., & Carnine, D. Reducing Behavior Problems: An Operant Guide for Teachers. Urbana, Illinois: Educational Resources Information Center, 1969.

Bialer, I. Conceptualization of success and failure in mentally retarded and normal children. Journal of Personality, 1961, 29, 303-320.

Butterfield, E. C. Locus of control, test anxiety, reactions to frustration, and achievement attitudes. Journal of Personality, 1964, 32, 298-311.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. Equality of Educational Opportunity. Washington, D.C.: U.S. Government Printing Office, 1966.

Crandall, V. Discussant's paper. Symposium on Developmental Aspects of Locus of Control Expectancies: New Methods and Prospects. Presented at the American Psychological Association Convention, Washington, 1971.

Crandall, V. C., Katkovsky, W., & Crandall, V. J. Children's beliefs in their own control of reinforcements in intellectual-academic achievement situations. Child Development, 1965, 36, 91-109.

Crandall, V. J., Katkovsky, W., & Preston, A. Motivational and ability determinants of young children's intellectual achievement behaviors. Child Development, 1962, 33, 643-661.

DeCharms, R. Personal Causation: The Internal Affective Determinants of Behavior. New York: Academic Press, 1968.

Franklin, R. D. Youths' expectancies about internal versus external control of reinforcement related to N variables. Unpublished doctoral dissertation, Purdue University, 1963.

Gruen, G. Response style. In W. Asher (Ed.), Final report of Phase I, USOE Contract No. OEC-0-70-4952: The development of new measures of cognitive variables in elementary school children. Purdue University, 1970.

Hudgens, G. A., & MacNeil, D. A. Aggressiveness and learning ability: Effects of histories of wins or defeats on avoidance learning in mice. Psychonomic Science, 1970, 20, 51-53.

Jessor, R., & Hammond, K. Construct validity and the Taylor anxiety scale. Psychological Bulletin, 1957, 54, 161-170.

Joe, V. C. Review of the internal-external control construct as a personality variable. Psychological Reports, Monograph Supplement 3-V28, 1971, 619-640.

Kelley, H. The processes of causal attribution. American Psychologist, 1973, 28, 107-128.

Lefconurst, H.M. Risk-taking in Negro and white adults. Journal of Personality and Social Psychology, 1965, 2, 765-770.

Levine, M. D., & Gordon, T. P. Emotional response of rhesus monkeys to chronic psychological stress. Proceedings of the Army Science Conference, 1968, 2, 17-25.

Logan, F. A, Incentive: How the Conditions of Reinforcement Affect the Performance of Rats. New Haven, Connecticut: Yale University Press, 1960.

Maier, S. F., Seligman, M.E.P., & Solomon, R. L. Pavlovian fear conditioning and learned helplessness. In Campbell and Church (Eds.), Punishment and Aversive Behavior. New York: Appleton-Century-Crofts, 1969.

McClelland, D. C., Atkinson, J.W., Clark, R. A., & Lowell, E. L. The Achievement Motive. New York: Appleton-Century-Crofts, 1953.

Mischel, W. Personality and Assessment. New York: Wiley, 1968.

Muller, J. A convergent assessment of children's locus of control measures: Th theoretical and methodological implications of divergence. Unpublished paper, Purdue University, 1972.

Nowicki, S., & Strickland, B. A locus of control scale for children. Journal of Consulting and Clinical Psychology, 1973, 40, 148-154.

Parker, L. Locus of control expectancies and the effects of aggression on performance on a learning task. Unpublished Master's thesis, Purdue University, 1971.

Rotter, J. B. Social Learning and Clinical Psychology. New York: Prentice Hall, 1954.

Rotter, J. B. Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 1966, 80, 1 (Whole No. 609).

Shore, M. F., Milgram, N. A., & Malasky, C. The effectiveness of an enrichment program for disadvantaged young children. American Journal of Orthopsychiatry, 1971, 41, 442-449.

Siegal, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.

Spence, K. W. Behavior Theory and Learning. New York: Prentice-Hall, 1960.

Stephens, M. W. Relation of achievement motivation to academic achievement in students of superior ability: Further data. Unpublished manuscript, Purdue University, 1969.

Stephens, M. W. Final report, USOE Contract No. OEC-0-70-4952, Phase III (Amendment): Evaluation of locus of control measures for Follow Through Evaluation. Purdue University, 1971 (a)

Stephens, M. W. Cognitive and cultural determinants of early IE development. Proceedings, 79th Annual Convention, American Psychological Association, 1971, 822. (b)

Stephens, M. W. NIMH Progress Report, Purdue University, 1972. (a)

Stephens, M. W. Dimensions of locus of control: Impact of early educational experiences. Proceedings, 80th Annual Convention, American Psychological Association, 1972, 137-138. (b)

Stephens, M. W., Delys, P., Lopez-Roig, L., & Vilez, C. Cultural factors in locus of control dispositions of Mexican-American and Puerto Rican children. Paper presented at InterAmerican Congress of Psychology, Panama City, 1971.

Stephens, M. W., Delys, P., & Parker, L. Cognitive, economic, and teacher determinants of locus of control expectancies. Unpublished paper, Purdue University, 1971.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

APPENDIX

SCORING MANUAL

General Scoring Rules

1. Score internal for response indicating reinforcement is contingent on the behavior of the subject.

2. Score internal for response for "self," "kids," "boys," or "girls" of same sex as subject.

3. Score internal for responses which indicate knowledge of a general classroom or social rule.

4. Score external for responses indicating generalized affect toward the child without stated cause.

5. Score external for responses which are internal on the part of someone other than the subject.

Scoring Examples
Form A

A. 1. What makes you happy?

Internal

Doing good things.
Riding my bike.
Myself.

External

When people give me candy.
God.
Mother.

2. What makes other kids happy?

Internal

When I play with them.
When I let them ride my bike.
When I give them cookies.

External

Doing things their mother tells them to do.
New toys.
When they do good things.

Special cases

Score internal for responses regarding others which could include the subject.

Examples: Other kids.
When their friends play with them.
When you are nice to them.

Note: Responses to this item must be internal on the part of the subject, not other children.

3. What makes mothers happy?

Internal

If you go to the store for her.
If you clean your room.
When kids are quiet.

External

When they get new clothes.
When she buys me a bike.
Daddy.

4. What makes fathers happy?

Internal

When little girls and boys don't fight.
When kids love him.
When I be quiet.

External

To go to work.
Mother.
To get new clothes.

Note: Score external for responses showing affect toward child without specific cause.

Examples: He likes children.
He likes to play ball with us.
He is not mad.

5. What makes teachers happy?
Internal
When kids sit in the circle.
When we be good.
When kids are quiet.
External
Clothes.
When she takes us on trips.
God and preachers.

B. 1. What makes you unhappy?
Internal
When I fight.
When I get my clothes dirty.
External
When someone hits me.
When I can't go anywhere.
Spiders and frogs.
Special cases
Score internal: When I'm mad at someone.

2. What makes other kids unhappy?
Internal
When someone hits them with a rock.
If I beat them up.
I do.
External
When they can't play in the dirt.
Their mothers and fathers.
When they get whipped.
Special cases
Score internal for behavior type responses which reflect knowledge of general classroom or social rules.
Examples: When they fight.
When they scratch someone.

3. What makes mothers unhappy?
Internal
When you don't be good.
When little kids act up.
Little kids.
External
When daddy won't take her along.
When she has to cook breakfast.
When rats come in the house.

4. What makes fathers unhappy?
Internal
Kids.
If you tell Mother he went out.
Kids screaming.
External
To have to go to work.
When he has to take Mama and the kids with him fishing.
Mother.

5. What makes teachers unhappy?
Internal
If kids play when they should work.
When kids be bad.
If you mess up the school.
External
To be mad.
When Randall always worries her.
She doesn't like kids.

C. 1. What makes you be nice?
Internal
Petting my puppy and cat.
When I share.
When we play at school.
External
Teacher.
When mother takes me to the store.
Food.

2. What makes other kids be nice?
Internal
Friends.
When kids let them play with their toys.
When I give them candy.
External
Riding bikes.
Teacher.
Food.

3. What makes mothers be nice?
Internal
When kids help her.
When the teacher tells her the kids are being good.
When you obey her.
External
When the police caught the bad man.
Father.
When Father is happy.

4. What makes fathers be nice?
Internal
When I'm a good boy.
When I be nice.
When Mother tells him the kids were good.
External
Playing basketball with his friends.
When he comes home.
To go to work.

5. What makes teachers be nice?
Internal
When I raise my hand.
When you be good.
When kids help her.

D. 1. What makes you be not nice?
Internal
When I don't like somebody.
Fighting.
When you don't let someone play with your things.
External
When Daddy whips me.
When Mother tells me to get out of the kitchen.
Kids. (Can't include subject)
Special case
Score internal for aggressive feelings.
Example: When I'm mad at somebody.

2. What makes other kids not nice?
Internal
When other kids don't be nice to them.
If you hurt them.
External
When they have to go to bed.
When they're bad.
If they have to do work.
Note: Responses must be internal on the part of the subject.
Special cases
General classroom or social rules.
Example: If they throw sand on the teacher.

3. What makes mothers be not nice?
Internal
Little kids.
When we don't wash our hands for dinner.
When kids be bad.
External
When they have to go to work.
The baby crying.
Because a gerbil bit me.

4. What makes fathers be not nice?
Internal
Little kids.
When kids disobey.
Kids screaming.
External
He's just mean.
When they're mad at people.
Headaches.

5. What makes teachers be not nice?
Internal
Little kids.
When we don't love her.
If we don't feed the gerbils.
External
When she is sick.
My mama because she got mad at the teacher.
God.

Form B

A. 1. What makes you smile?
Internal
When I'm busy.
When I be good.
External
When someone gives you something.
The sun.
When people tickle me.

2. What makes other kids smile?
Internal
If you buy them something.
If you share with them.
If I play with them.
External
When they obey.
When they get new clothes.
When they eat lunch.

3. What makes mothers smile?
Internal
When she comes home and you have been good.
When you buy them presents.
Kids.
External
When Daddy comes home.
When she's happy.
When Daddy kisses her.

4. What makes fathers smile?
Internal
When kids be good.
When you do something for him.
When I sit on his lap.
External
When he's happy to see his friends.
When he has lots of money.
When he can go to work.

5. What makes teachers look mean?
Internal
When kids act bad.
When we don't act right.
When I be noisy.
External
When the dog bites her.
When the man didn't bring the food.
When she's mad.

C. 1. What makes you feel good?
Internal
When I'm good.
Because I like boys and girls.
If I share.
External
When we go to the airport.
Mama.
Teacher.

2. What makes other kids like you?
Internal
Because I play with them.
When I'm their friend.
If you share.
External
They don't like me.
Because they are happy.
Because we can take our toys out.

3. What makes mother love you?
Internal
When I go to school.
Me.
When I'm real happy.
External
When she sends you outdoors.
Because she likes us.
Because she's happy.

4. What makes father love you?
Internal
When you be good.
When you're happy and smile.
When you go to bed when he says to.
External
Because he likes me.
Because he's happy.
Because he likes little people.

5. What makes teachers smile?
Internal
When kids be quiet.
Good kids.
When you do something right.
External
When things are funny.
The sun.
When my Daddy comes to talk to her.

B. 1. What makes you look mean?
Internal
When I mess up when I write.
When I'm going to fight someone.
If I don't share.
External
If kids fight me.
My mother.
When I can't go outdoors.

2. What makes other kids look mean?
Internal
If I hit them.
When you close the door on them.
If someone hits them.
External
If they cry.
The devil.
When they can't go out.

3. What makes mothers look mean?
Internal
Kids.
If you fight.
If you don't shut up.
External
If father beats her up.
If the baby spills his food.
When she falls in the river.

4. What makes fathers look mean?
Internal
If I hit him.
When someone messes with him.
If kids eat with dirty hands.
If we be bad.
External
If the baby jumps on him.
When Mama yells at him.
When he paints.

5. What makes teacher like you?
Internal
Because I'm a good girl.
When we play good.
When I beat the boy up.

D. 1. What makes you angry?
Internal
Myself.
When mama yells at me for getting muddy.
When I cry.
External
When I get whipped.
Little kids. (Can't include subject)
No food.

2. What makes other kids angry?
Internal
If people hit them.
If you tell on them.
If you fight with them.
External
When they have to go in the other room.
Because they can't play.
When they have to take their clothes off.

3. What makes mothers be angry?
Internal
Kids.
If kids run away.
If you do something bad.
External
My brother.
To have another baby.
Daddy.

4. What makes fathers angry?
Internal
Little kids.
When kids be bad.
When he tells you to do something and you don't.
External
When he can't eat.
When he fusses at the kids.
Because he don't like police.

5. What makes teachers mad?
Internal
When I don't be good.
If we fight.
External
If she's not happy.
Because she doesn't like boys.
She's a lady.