Environmental Input and Cognitive Growth: A Study Using Time-Period Comparisons

Janellen Huttenlocher, Susan Levine, and Jack Vevea

In this study, we examined the relation of input to cognitive growth in a single population of children. We studied 4 domains: Language, Spatial Operations, Concepts, and Associative Memory. Four groups of children drawn from the same population were tested in October of kindergarten, April of kindergarten, October of first grade, and April of first grade. These time points are 6 months apart, but they span periods that differ in amount of school input children receive. Much greater growth was found over time periods with greater amounts of school input (October to April) than over time periods with less school input (April to October) for Language, Spatial Operations, and Concepts, but not for Associative Memory. These findings suggest that amount of input is causally related to cognitive growth in particular domains.

INTRODUCTION

Currently there are differing views about the role of environmental input in cognitive growth. One widely held view stresses the plasticity of the developing brain/cognitive system, focusing on the role of input in the cognitive skill levels that individuals achieve. Many studies of cognitive development show that variation in input is related to achieved skill levels. First, children from homes providing more input do better on tests of ability and tests of achievement (e.g., A. W. Gottfried, 1984; Honig, 1982; Kellaghan, Sloane, Alvarez, & Bloom, 1993), and language development is significantly related to amount of vocal stimulation at home (e.g., Clarke-Stewart, 1973; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1992; Snow & Ferguson, 1977). Second, children from schools providing more input attain higher skill levels than children from schools providing less input. For example, Chinese children, who receive more mathematics instruction at school than American children, attain higher levels of mathematical skill than American children (e.g., Stevenson, Lee, & Stigler, 1986; Stigler, Lee, & Stevenson, 1987). Animal studies also show that variation in environmental input is related to brain development as indexed by the thickness of the cerebral cortex, dendritic elaboration, and so forth (e.g., Diamond, Krech, & Rosenzweig, 1964; Greenough & Volkmar, 1973).

An alternative view, widely held by behavior geneticists, stresses differences in intellectual potential (ability) across people, focusing on the role of hereditary factors in the cognitive skill levels that individuals achieve (e.g., Plomin & DeFries, 1980; Scarr, 1992). Studies investigating the role of heredity in children's skill levels generally measure their intellectual potential through the scores their biological relatives attain on intelligence tests. The measure of children's skill levels is generally their own scores on intelligence tests. Many studies show that children whose biological parents have higher IQ test scores themselves have higher IQ test scores. In the view of these investigators, genetic factors severely constrain the plasticity of the developing brain/cognitive system.

Designs for Studying the Sources of Individual Differences

Problems of interpretation arise for many studies attempting to identify the sources of individual differences in cognitive skill because these studies are correlational. Direction of cause is ambiguous, and observed relations may be mediated by factors other than those which the study identifies and examines. Because intellectual potential (parent IQ) and input variables tend to covary in natural environments, they are likely to be confounded in studies of the relation between individuals' cognitive skill levels and one of these variables (intellectual potential or amount of input). These problems are especially great because both the measures of the child's potential and the measures of input to the child are themselves problematic.

Assessing the child's potential using IQ of biological relatives is problematic because IQ is not free of possible influences of input. The lower IQs of children from families with lower IQs may reflect, at

least in part, less adequate input at home and/or at school. In reviewing the literature on the relation between IQ and amount of schooling, Ceci (1991) argues that, contrary to the traditional view that IQ affects the amount of schooling people receive, there is strong evidence suggesting that amount of schooling affects the IQ scores people obtain.

Using amount of input to assess environmental factors is problematic as well because input is not free of possible influences of intellectual potential. Families that provide more adequate input at home and live in neighborhoods with better schools, and so forth, may do so because they have higher ability levels. Behavior geneticists also argue that, to some extent, children determine their own input, choosing different inputs according to their abilities (see Scarr, 1992).

Given these problems of design and measurement, it is not surprising that, despite wide agreement that achieved skill levels reflect an interaction between environmental input and intellectual potential, there are differences among investigators in the relative importance they attribute to each of these factors. To determine definitively that there are substantial effects of input or of intellectual potential on achieved skill levels, it clearly would be advantageous to use designs where these two factors do not covary. A design widely used by behavior geneticists to examine variation in potential that is uncorrelated with input involves examining children who are not reared by their biological families (see Plomin & DeFries, 1980; Scarr, 1992). Adoption studies in which identical twins are raised in different families provide the clearest control of hereditary factors because their genetic characteristics are completely shared (Bouchard et al., 1990). Although questions have been raised about the conclusions that can be drawn from these studies (e.g., Kamin, 1974), even skeptics generally agree that the studies show at least some causal role of heredity in the skill levels individuals achieve.

Investigators concerned with input effects on cognitive skill have not so explicitly articulated designs for effectively studying cognitive skills under naturally occurring conditions in which input and potential do not covary. Studies of school effects can potentially provide a way to investigate input effects. Because teachers are not biological relatives of the children studied, relations between school input and achievement cannot be directly linked to hereditary similarities between children and input providers. However, simple correlational studies of the relation between schooling and level of cognitive skill are still problematic, because the quality of schools is related to the characteristics of the population groups they serve. These difficulties can be avoided if planned interventions are introduced, with random assignment of children to intervention and control conditions. Extensive interventions involving such designs have been used in studies of impoverished preschool children. Children in the intervention condition showed higher IQ scores in the preschool period (Ramey & Campbell, 1984), and, for verbal scores, these children still had higher scores at 12 and 15 years (e.g., Campbell & Ramey, 1994, 1995).

Two designs have been introduced in an attempt to explore naturally occurring variations in school input under conditions in which input and ability do not covary. We refer to these as the age cutoff and time-period comparison designs. Insofar as these designs successfully control for ability levels, it would be possible to evaluate the role of input in the growth of cognitive skills which are commonly believed to reflect the intellectual potential of individuals (i.e., ability). Thus far these designs have been used primarily to examine the growth of curriculum-related skills rather than of skills that are believed to reflect primarily intellectual potential.

The age cutoff design was introduced by Baltes and Reinert (1969) to study input effects with children of similar age whose school experiences vary because of the fact that schools use age cutoffs in making grade placements. If the age cutoff is used without exception, it would be possible to compare groups of children whose birthdays fall either just before or just after the cutoff. These children would vary little in age or intellectual potential, but would differ in the amount of school input they have received. This design was used by Morrison and his colleagues (e.g., Bisanz, Morrison, & Dunn, 1995; Morrison, Smith, & Dow-Ehrensberger, 1995) to study the cognitive skills of kindergarten and first-grade children. Morrison et al. (1995) found that scores on reading-related tasks (phonological awareness, knowledge of sound-letter correspondences, and word decoding) were related to grade, not age, and, similarly, Bisanz et al. (1995) found that number fact tasks were related to grade, not age. Grade should affect performance on reading and number fact tasks because children in the higher grade are exposed to a more advanced curriculum. If successively more complex materials are presented each year, one would expect a substantial grade effect on skill.

In addition to effects of more advanced curricula, children in a higher grade have had an extra year of schooling. Thus if the input relevant to cognitive growth is greater at school than at home, then there

should be at least some grade effect even if that input does not form part of a successively advancing curriculum. For this reason, if the groups are equal in intellectual potential, the age cutoff design could be used to determine if input affects performance levels on skills believed to reflect differences in intellectual potential (ability). However, there is some evidence that the groups may not be equal; for example, A. W. Gottfried, A. E. Gottfried, Bathurst, & Guerin (1994) found that gifted children tended to be younger at kindergarten entry than their comparison cohort. Hence comparison of children of similar age in different grades may not provide a perfect design for examining input unconfounded with ability.

The time-period comparison design examines input effects in a single population group, thus avoiding the possibility that intellectual potential covaries with input. Growth is compared across time periods that vary in amount of schooling—the school year versus the summer. In a study using the time-period comparison design, Heyns (1978) examined vocabulary growth in fifth to seventh grade children. She found much greater gains over the school year than over the summer. Entwistle and Alexander (1992) have used the design to study mathematics achievement in the first 2 years of school. Their findings indicate that there is greater growth of mathematical skill over the school year than the summer. The findings provide evidence that change over the summer varies with socioeconomic group. That is, children from higher socioeconomic groups show larger gains or smaller losses than children from lower socioeconomic groups (Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996).

An uneven pattern of growth in the same individuals or population groups associated with time periods that vary in input can provide unequivocal evidence of the input sensitivity of a skill. Thus the design can be used to determine if input affects performance levels on skills believed to reflect differences in intellectual potential (ability). Note, however, that because the method relies on differential growth, the input sensitivity of a skill will be detected only if the inputs critical to growth are more available at school than at home. As we have noted, middle-class children may show gains over the summer as well as the school year. Hence it would be possible that middle-class children might not show differential growth during the school year versus the summer even for skills that are input sensitive. However, insofar as summer gains are smaller than school year gains, as suggested in the meta-analysis by Cooper et al., time-period comparisons can reveal input sensitivity even in middle-class children.

Assessment of Cognitive Skills

The manner in which investigators classify the cognitive skills people exhibit may substantially affect the conclusions drawn about the sources of individual differences. However, there is presently no single generally accepted way to classify cognitive skills. One well-known theoretical distinction is clearly related to the issue of input sensitivity—the distinction between "fluid intelligence" that involves mainly intellectual potential and "crystallized intelligence" that involves particular achievements and thus necessarily reflects effects of input as well as potential (see Guilford, 1967).

The notion that tasks may differ in the extent to which they draw on intellectual potential versus learned skills is widely recognized in the literature on standardized tests; there is a general division between tests designed to assess ability and those designed to assess achievement. Ability tests are designed to tap relatively stable cognitive processes that are not explicitly taught and to assess mainly an individual's potential. Achievement tests are designed to tap input-driven skills that require extensive learning and generally are part of the school curriculum; for example, reading and arithmetic. If ability and achievement tests measure distinct aspects of intellectual functioning, they could be used to study the relation between achievement and ability in the levels of skill people attain. However, these two kinds of tests actually overlap markedly with respect to the tasks included. For example, both types of tests include vocabulary and general information scales, and for both, scores on vocabulary scales are highly correlated with overall scores, with correlations ranging from .71 to .98 (Anderson & Freebody, 1981).

An alternative way to categorize cognitive skills is according to domain. The most widely accepted distinction is that between verbal and spatial domains, although larger numbers of domains have been suggested by others (e.g., Gardner, 1983; Thurstone, 1938). There are several bases for drawing this distinction between verbal and spatial domains. Correlations between verbal and spatial skill levels in normal adults tend to be relatively low (e.g., Thurstone, 1938); levels of skill in these domains are distributed differently in different population groups (notably between the sexes, cf. Maccoby & Jacklin, 1974), and there are differential effects of damage to particular areas of brain on language and spatial skill levels (e.g., Sperry, Gazzaniga, & Bogen, 1969; Teuber, 1955). It should be noted that not just investigators concerned with individual differences, but

also those concerned with commonalities in cognitive development, have posited that distinct processes are involved in different skills and that these constitute separate domains that may be supported by different brain structures (e.g., Karmiloff-Smith, 1992).

If there are separable domains of skill, it is possible that these domains may vary in the extent to which performance levels are input driven. For example, language growth might be more (or less) sensitive to input than spatial growth. Even if the extent to which two domains are input dependent is equal, the type of input relevant to growth in those domains may differ. For example, amount of talk is the kind of input important for language skill, and amount of puzzle and construction activity may be important for spatial skill. An individual child might receive more input relevant to one domain than another, leading to a higher skill level in that domain. If distinct cognitive domains exist, it would not be possible to properly evaluate the role of input in cognitive growth without assessing each of them separately.

Although the possibility that there are distinct cognitive domains is important to exploring the sources of individual differences in young children, the issue has not been systematically explored. Most standardized ability tests for children, such as the Stanford Binet Intelligence Scale (Thorndike, Hagen, & Sattler, 1986), and the Wechsler Intelligence Scale for Children (Wechsler, 1989), are designed to evaluate overall ability levels rather than to determine if distinct domains of cognitive skill exist. Many of the tasks included in these tests require a wide range of skills, so a child's level of function in one domain may be obscured by a lower level of function in another. This alone could produce a substantial correlation among different tasks, making it appear that there is just one general cognitive domain even when there are several separate domains.

In our earlier work we developed a set of tasks designed to identify distinct domains of cognitive skill in kindergarten and first-grade children. The set of tasks was analyzed to determine whether separate domains could be identified and what these domains would consist of. The investigation led to the construction of a test (The Primary Test of Cognitive Skills, PTCS; Huttenlocher & Levine, 1990). This test is used in the present study to examine the effects of input on cognitive growth. In developing the test, our approach to the identification of distinct domains was to begin with the question of whether there are distinct language and spatial domains in young children as there are in adults. We developed a set of tasks that might potentially tap language skill and a set of tasks that might potentially tap spatial skill.

Analysis of the tasks revealed four distinct domains. We describe the test and its development in the Appendix because, as indicated above, the establishment of what constitutes distinct and basic cognitive skills is critical to evaluation of input sensitivity.

The Present Study

The present study was designed to examine the extent of input sensitivity in the growth of basic cognitive skills. We used the test we developed, which showed four distinct domains of skill, to examine the relation of school to growth in each domain. The study focused on the question of whether we would find school effects for the Language and Spatial Scales, because these skills are commonly regarded as reflecting intellectual potential rather than achievement. We also obtained data on the other two scales (Concepts and Memory Scales). We expected to find school effects for the Concepts Scale, because these items assess, in a nonverbal format, conceptual knowledge that might be encompassed in the school curriculum. We had no strong expectations as to whether the Memory scale would show school effects.

The input sensitivity of skill levels in these cognitive domains was investigated using the time-comparison design. Children were tested at four time points that were equally spaced, spanning periods that differed in the extent of schooling. This made it possible to compare growth in kindergarten and first-grade children from a wide range of backgrounds over time periods which differed in input. The study was cross-sectional, that is, different children were included in the different test groups. However, because the children in the different test groups were drawn from the same population group and the sample was very large, the study is quasi-longitudinal. An uneven pattern of growth indicates input sensitivity for the population, although not for individuals.

As noted above, the time-period comparison design has not thus far been used to systematically investigate whether skills commonly believed to reflect primarily intellectual potential, such as language and spatial transformation, are input sensitive. An uneven pattern of growth for a skill, related to the school year, would provide strong evidence of input as a factor. One might expect that tasks which combine to form a single scale—for example, the syntactic and lexical aspects of language—might be sensitive to the same types of input. If, instead, the component tasks were sensitive to very different aspects of input, that could well detract from the corre-

lation between those tasks because those different aspects of input might vary independently of one another.

METHOD

Design and Procedure

Children were tested at four time points (October of kindergarten, April of kindergarten, October of first grade, and April of first grade). These contrasting periods are of equal length, but they vary markedly in the extent of input from schooling. That is, from October to April, children are in school, whereas from April to October there is summer break, as well as the time when work is getting started at the beginning of the school year and winding down at the end of the school year. Although there is some schooling over the April to October period, it is considerably less than during the October to April period. As noted above, our design was cross-sectional. Four large groups were tested during a single school year, two in the fall and two in the spring. The differences in performance among the groups provide a measure of growth over the population during different time periods.

Children were tested in small groups in their classroom and were presented the items by their teachers. All testers followed a standardized set of instructions as presented in the PTCS manual (Huttenlocher & Levine, 1990). Children indicated their answers by filling in "bubbles" under the proper choice in booklets. This procedure was taught during two practice sessions where children learned to fill in bubbles and to find the item on the page that the teacher presented each time. In addition, children were given items representative of the various item types in all scales during the practice sessions. The Language Scale consisted of 15 vocabulary and 9 syntax items, the Spatial Operations Scale consisted of 10 sequencing and 11 integration items, the Concepts Scale consisted of 13 object and 15 spatial items, and the Associative Memory Scale consisted of 7 sound-object and 14 object-location items. Reliabilities (Cronbach's alpha) for the language, spatial operations, concepts, and memory scales are .80, .81, .84, and .72, respectively.

Participants

Participants consisted of a national sample of kindergarten and first-grade children from randomly selected schools. Sampling cells were based on type of school (public, private, and parochial), geographic sections of the country (northeast, midwest, south-

east, and west), community type (urban, suburban, and rural), and community socioeconomic level as indicated by percentage of children receiving Chapter 1 funding (Orshansky, 1988a, 1988b). All four groups (Fall Kindergarten, Spring Kindergarten, Fall First Grade, and Spring First Grade) were selected in the autumn of a single academic year and were tested during that year. The selection of schools in each group followed the usual procedures of the California Test Bureau to obtain a representative sample mirroring the population of various groups in the American population. Hence the four groups can be regarded as equivalent. There was oversampling to ensure adequate sample size given that there typically is attrition, especially in the spring, due to competing commitments at the schools. There were 2,033 kindergartners tested in October of the school year, 1,652 kindergartners tested in April of the school year, 2,387 first graders tested in October of the school year, and 1,790 first graders tested in April of the school year. The number of males and females in each group was approximately equal.

RESULTS

Growth over Time

To examine growth over time in such a way that the four cognitive domains were comparable, we calculated a z score for each child on each scale. The z scores were produced using a single marginal mean and standard deviation for each scale, derived from the present sample without regard to age or sex. The effect of that transformation is to make the units for each scale have the same meaning, so that any observed changes are readily interpretable. Figure 1 shows the pattern of means across ages by sex for Language, Concepts, Spatial Operations, and Associative Memory, respectively. It is worth noting that, for any particular time of measurement, the absolute magnitudes of sex differences were much smaller than differences observed across time for a particular sex. Sex differences were on the order of less than one-tenth of a standard deviation (average absolute difference = 0.07, maximum = 0.17), whereas time differences within a level of sex could be as high as three-quarters of a standard deviation. Thus, the practical significance of sex is very small compared to that of time.

Table 1 shows the average percent correct for each group on each scale, as well as the mean of the corresponding z score; standard deviations for the z scores appear below the means. The lower part of the table gives the same marginal means and standard devia-

tions for the two sexes. Table 2 shows the means and standard deviations within age × sex cells for the Language, Concepts, Spatial Operations, and Memory scales, respectively. We performed a multivariate analysis of variance (MANOVA) incorporating tests for a time effect, a sex effect, and an interaction. Both main effects were highly significant. For sex, Wilks's lambda was 1.00, F approximation(4, 7876) = 8.15, $p < .001$. For time, Wilks's lambda was 0.66, F approximation(12, 20838.2) = 301.45, $p < .001$. However, the multivariate test for an age × sex interaction was marginal, Wilks's lambda = 1.00, F approximation(12, 20838.2) = 1.66, $p = .07$.

We followed up the multivariate analysis with univariate tests of the sex and time of measurement main effects. Note that, although the interaction was borderline in the multivariate test, that near-significance is primarily the consequence of the great power associated with our large sample size; thus, we elected not to investigate the univariate interactions. The univariate tests for sex effects were: for memory, $F(1, 7879) = 31.61, p < .001$, for spatial operations, $F(1, 7879) = 5.00, p < .03$, for concepts, $F(1, 7879) = 5.04, p < .025$; and for language, $F(1, 7879) = 4.69, p < .05$. The tests for time effects were: for memory, $F(3, 7879) = 487.72, p < .001$, for spatial operations, $F(3, 7879) = 2133.64, p < .001$, for concepts, $F(3, 7879) = 866.02, p < .001$, and for language, $F(3, 7879) = 523.86, p < .001$. As we will discuss below, summer growth was considerably less than growth over the school years for these scales. Nevertheless, Bonferroni adjusted post hoc comparisons revealed that all possible pairwise juxtapositions of means, even the means spanning the summer period, were significant at an overall alpha of .05.

Growth during School versus Summer

Table 1 shows clearly that growth for Language, Concepts, and Spatial Operations was smaller over the summer than over each of the school years. For Memory, in contrast, there was no such effect; growth rate simply decreased with time. Figure 2 shows the cumulative percent of children receiving scores up to particular levels in each group—Fall K, Spring K, Fall 1, Spring 1. Each graph for a domain has four curves, one for each of the time points, showing the differential growth over the school years and over the summer. The school year–summer contrast can be seen in the greater distances between the first two curves and between the last two curves than between the central two curves for the three scales that show a school effect. The fact that there was not a school effect for Memory serves as a natural control for the possibility that greater growth of the school year might be artifactual, reflecting greater practice in concentration, filling in "bubble sheets," and so forth.

To examine growth in school versus summer statistically, we take the average of the two school year growth periods and contrast this average growth to summer growth. That is, because we have two opportunities to measure school year change, once during kindergarten and once during first grade, we take the average of these two changes and compare this with the change over summer. The contrast representing average school year growth is $(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2})$. The contrast representing summer growth is $(0, 1, -1, 0)$. Subtracting the first from the second produces the coefficients $(-\frac{1}{2}, \frac{3}{2}, -\frac{3}{2}, \frac{1}{2})$, which may be conveniently rescaled to $(-1, 3, -3, 1)$. Thus the coefficients $(-1, 3, -3, 1)$ compare summer growth and average school year growth. Table 3 presents a point estimate and 95% confidence interval for the contrast on each outcome. Because the scales are expressed in the z score metric, the contrasts can be interpreted as differences between summer and school year growth in standard deviation units. The contrast for Spatial Operations is associated with an $F(1, 7879) = 36.45, p < .001$, for Concepts, $F(1, 7879) = 77.20, p < .001$, and for Language, $F(1, 7879) = 71.65, p < .001$. For Memory, on the other hand, $F(1, 7879) = 0.99, p = .32$. For Language and Concepts, growth over the summer is about three-quarters of a standard deviation lower than growth over the school year. For Spatial Operations, growth over the summer is about one-half of a standard deviation lower than growth over the school year. The difference is near zero for Memory.

Let us consider whether the differences between growth over the school year and over the summer are the same among the three scales that show school effects—the Language, Concepts, and Spatial Operations scales. Within any one of those scales (Language, for example), the school-summer growth difference was represented by the contrast with coefficients $(-1, 3, -3, 1)$. The question of whether that contrast is the same for two scales (e.g., Language and Spatial Operations) may be addressed by applying the same contrast to a comparison between school-summer growth differences of the two scales (e.g., calculated by subtracting the standardized Spatial Operations score from the standardized Language score for each individual). When that comparison is made for the Language and Spatial Operations scales, the contrast estimate is 0.261, $F(1, 7879) = 7.11$, $p = .007$. Thus, there is highly reliable evidence that whereas both scales show improved school-year growth compared with summer growth, the magnitude of the school year–summer contrast is greater

A   Language
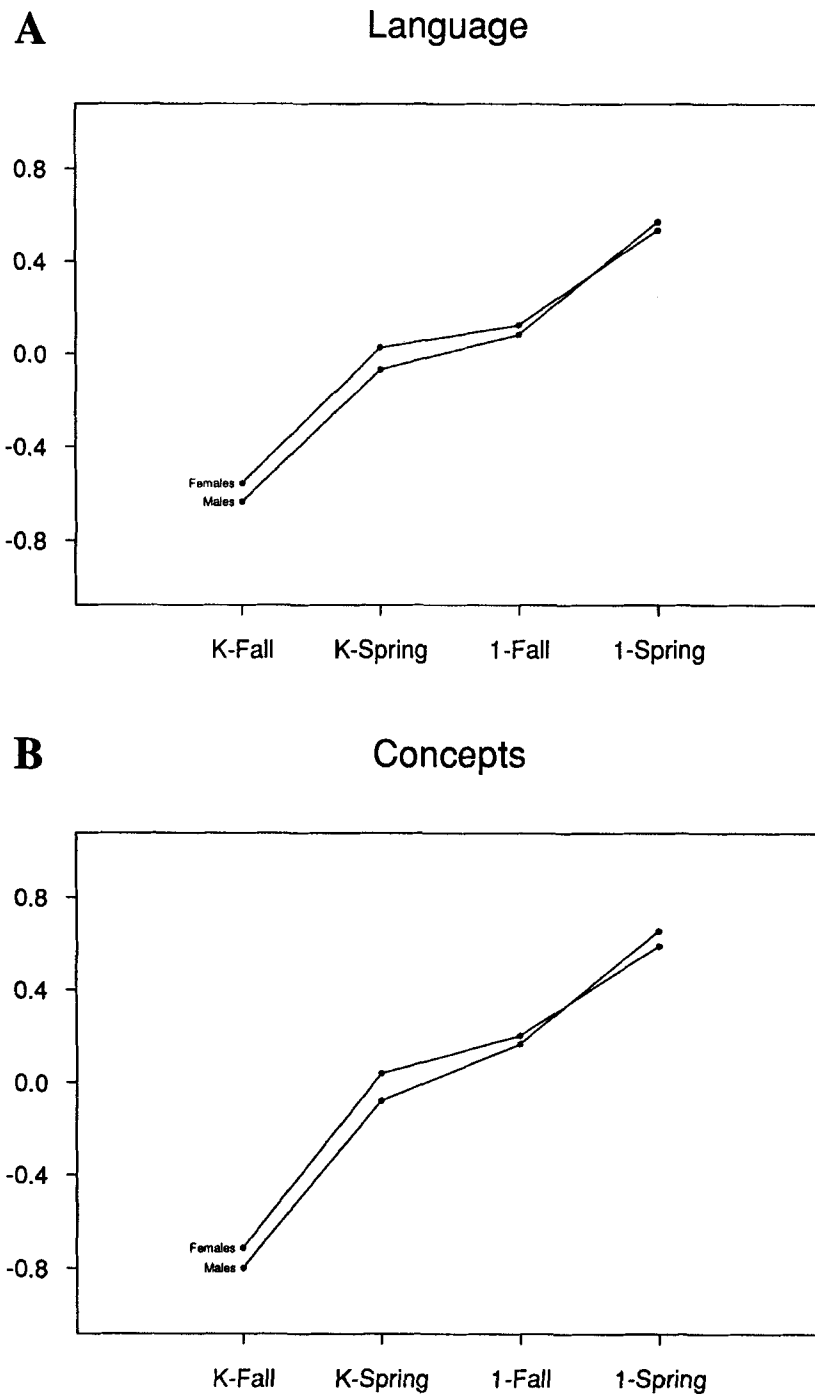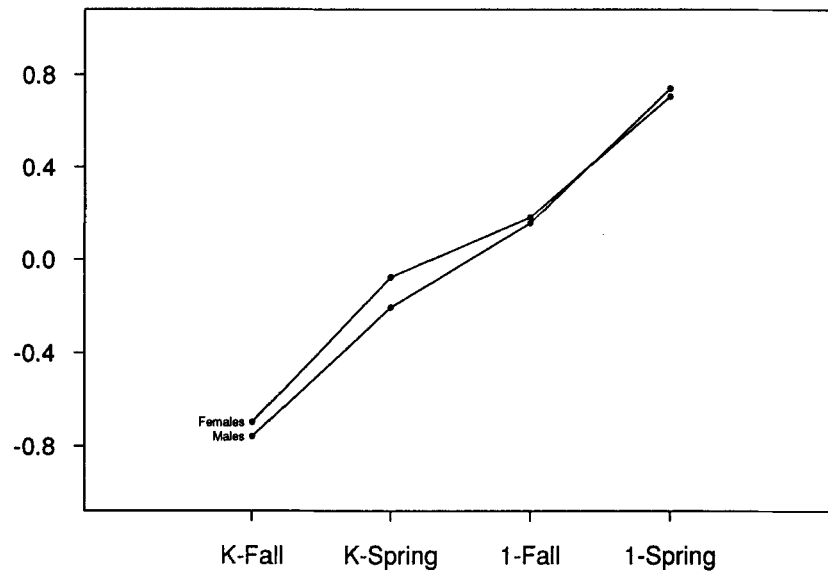

B   Concepts

Figure 1   Mean Z scores on each task for male and female children at each of the four times of testing for (A) Language, (B) Concepts, (C) Spatial Operations, and (D) Associative Memory (K = kindergarten, 1 = Grade 1).

by 0.26 standard deviation units for Language than for Spatial Operations.

When we compare Concepts and Spatial Operations, the contrast estimate is 0.25, $F(1, 7879) = 7.56$, $p = .006$. Concepts tasks, then, like Language tasks, show a greater school-year versus summer differen-

tial than do Spatial Operations. Finally, we find that the contrast that compares the school versus summer growth differences for the Language and Concepts scales is only -0.01, $F(1, 7879) = .01$, $p = .899$, ns. These results indicate that, whereas the Spatial Operations scale shows a substantially higher degree of

C   Spatial Operations

D   Associative Memory
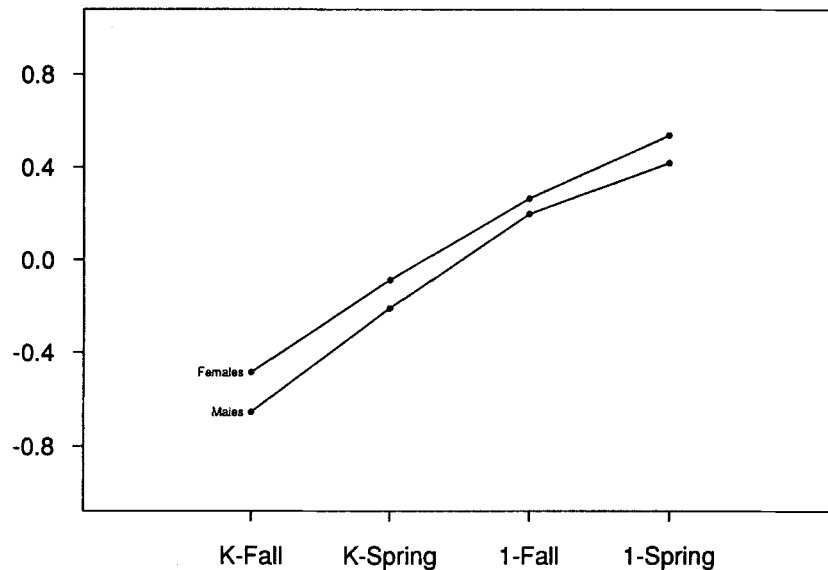
growth over the school year than over the summer, the Language and Concepts scales both show even greater differences between summer and school-year growth rates.

As we noted above, we have averaged across grades to make our statistical comparisons between summer and school year effects. There is some possibility that, in the case of memory, averaging might mask a school effect. Growth in memory showed a downward trend as age increased. If school effects are small when compared to that trend, combining estimates of school year growth over kindergarten and first grade might make it more difficult to detect such an effect. Therefore, let us consider the contrast of kindergarten year growth with summer growth separately from the contrast of first-grade growth

Table 1   Raw and Standardized Scale Scores for Each Age and Sex Group

	Language		Concepts		Spatial Operations		Memory	
	Percent Correct	Mean z Scores	Percent Correct	Mean z Scores	Percent Correct	Mean z Scores	Percent Correct	Mean z Scores
Fall K	51	−.597 (.916)	53	−.757 (.980)	45	−.729 (.793)	46	−.570 (.990)
Spring K	61	−.020 (.937)	67	−.018 (.930)	57	−.141 (.892)	53	−.148 (.927)
Fall 1	63	.103 (.905)	71	.184 (.822)	63	.169 (.883)	60	.230 (.874)
Spring 1	72	.555 (.899)	79	.625 (.721)	74	.726 (.846)	64	.474 (.885)
Male	61	−.142 (1.002)	67	−.012 (1.107)	59	−.011 (1.028)	55	−.048 (1.021)
Female	62	.015 (.997)	67	.013 (.981)	60	.012 (.969)	57	.051 (.974)

Note: Quantities in parentheses are standard deviations.

with summer growth. The contrast representing kindergarten year growth is (1, −1, 0, 0). The contrast representing summer growth is (0, 1, −1, 0). Subtracting the first from the second gives (−1, 2, −1, 0), which thus compares kindergarten year growth with summer growth. The estimated difference in growth for kindergarten and summer is 0.04, $F(1, 7879) = 0.65$, $p = .422$. First-grade growth may be presented by the coefficients (0, 0, 1, −1). Subtracting these from the coefficients for summer growth yields the contrast (0, 1, −2, 1), which compares first-grade growth with summer growth. The difference in growth is −0.14, $F(1, 7879) = 7.60$, $p = .006$. Thus, growth in associative memory was reliably higher over the summer than during the first-grade year, but the difference between growth in memory during kindergarten is not significantly higher than growth in memory over the summer. Thus the conclusion that

Table 2   Raw and Standardized Scale Scores for Each Age by Sex Group

	Fall K		Spring K		Fall 1		Spring 1	
	Percent Correct	Mean z Scores	Percent Correct	Mean z Scores	Percent Correct	Mean z Scores	Percent Correct	Mean z Scores
Language:								
Male	50	−.637 (.928)	60	−.067 (.891)	63	.083 (.932)	72	.572 (.868)
Female	52	−.558 (.902)	62	.028 (.979)	64	.125 (.874)	71	.536 (.933)
Concepts:								
Male	52	−.799 (1.006)	66	−.076 (.919)	70	.167 (.858)	79	.655 (.684)
Female	54	−.714 (.950)	68	.042 (.939)	71	.204 (.778)	78	.591 (.758)
Spatial operations:								
Male	45	−.760 (.811)	56	−.205 (.906)	63	.157 (.904)	74	.742 (.879)
Female	46	−.698 (.772)	58	−.075 (.874)	63	.183 (.859)	73	.707 (.808)
Memory:								
Male	44	−.654 (.995)	52	−.209 (.938)	60	.199 (.916)	63	.418 (.894)
Female	47	−.485 (.979)	54	−.086 (.911)	61	.265 (.823)	66	.5 (.871)

Note: Quantities in parentheses are standard deviations.

school effects for memory are not greater than home effects appears to be supported. The downward trend in growth in memory is not masking a major school effect.

In summary, performance on the Concepts and Language scales shows much larger differences in growth over the school year compared to growth over the summer. Although this is true of the Spatial Operations scale as well, the difference is less marked. In contrast, Memory shows no evidence of differential input effects between school and summer periods. Although one must always be cautious when interpreting negative results, the power of the present study, given the large number of participants, is such that very small effect sizes could be reliably detected; hence, it seems reasonable to place credence in the negative result for Memory.

Growth and starting level. Let us consider whether the effects of school are similar or different depending on the level of achieved competence at which a child begins. The results can be seen in Figure 1, from the cumulative curves at the four time points in the different domains. Note that these curves look roughly parallel. That is, the gain across a time period appears to be similar at all levels. To check this, we examined whether the school year–summer contrast for children at different ability levels statistically differs for the scales that show strong school effects on growth—the Language, Concepts, and Spatial Operations scales. For Language and Concepts, the size of the school effect did not differ significantly when the contrasts were considered for the different ability levels. That is, when attention was restricted to each quartile of each age group separately, the same trend was present. For each of the quartiles—first, second, third, and fourth—considered separately, the contrasts were significant at $p <$ .001. For Spatial Operations, however, the school effect was not constant; that is, for the bottom three quartiles, the same pattern holds; however, for the fourth quartile, the highest group, the contrast between growth during the school year and the summer is not significant.

A look within scales. Recall that test construction began with eight tasks which were then combined to form four scales. The fact that pairs of tasks formed unidimensional scales suggests that either both or neither of the two subtasks of a scale should show school effects. Otherwise, the subscales would not be sufficiently related to form a single scale. However, there is reason to pursue the issue of subscales further. For example, for language, as noted in the "Discussion" below, there are claims in the literature that syntactic growth is a modular ability, separate from

lexical growth. Because we obtained striking input effects on our Language scale, which includes both vocabulary and syntax items, it is desirable to check whether each subscale, looked at separately, shows school effects. At the same time this comparison can be made for the subscales for the other three scales as well. The results are shown in Table 4. Table 5 shows the significance of the growth differences over the school year versus those over the summer.

We checked dimensionality for just those items selected for the final scale using the present sample of participants. Note, however, that the scale was not built to separately sample the full range of item difficulty equally or to maintain high reliability when further subdivided. Hence, we calculated reliability coefficients (Cronbach's alpha) for the eight subscales. Although in some cases these shorter scales have low reliability, the sample size of the present study is such that quite stable conclusions about the population can be reached even with relatively unreliable measurements of individuals. Reliabilities for the vocabulary and syntax items of the Language scale were .77 and .51, respectively. The knowledge-based concepts items had a reliability of .66, and the spatial concepts items had a reliability of .83. Within the Spatial Operations scale, alpha was .77 for sequencing items and .66 for integration. Reliabilities for the Memory subscales were .71 for object location tasks and .44 for sound-picture association.

When we looked at just the items included in the final subscales to check dimensionality, we found weak evidence of slight dependence between the residuals within the separate tasks for two of the scales—Memory and Spatial scales. The evidence against unidimensionality even on these two scales was not strong. Nevertheless, as a check on our conclusions concerning school versus home effects for the four scales, we examined these effects for the two subparts of each. Table 4 shows the mean percent correct and z score for each of the subscales by group. The growth patterns exhibited by the means appear very similar to the patterns for the composite scales, with the exception of the Memory scale. There, the object location items seem to show the same pattern as the overall Memory scale, namely, relatively consistent growth over the school year and summer, and slower growth over first grade. The sound-picture items, however, seem to show the same pattern as the non-memory-related tasks: growth is higher during the school year and lower over the summer. An analysis of variance bears out those impressions.

The contrasts comparing school year versus summer growth are given in Table 5. The first six subscales show the same pattern as their parent scales:

A   Language

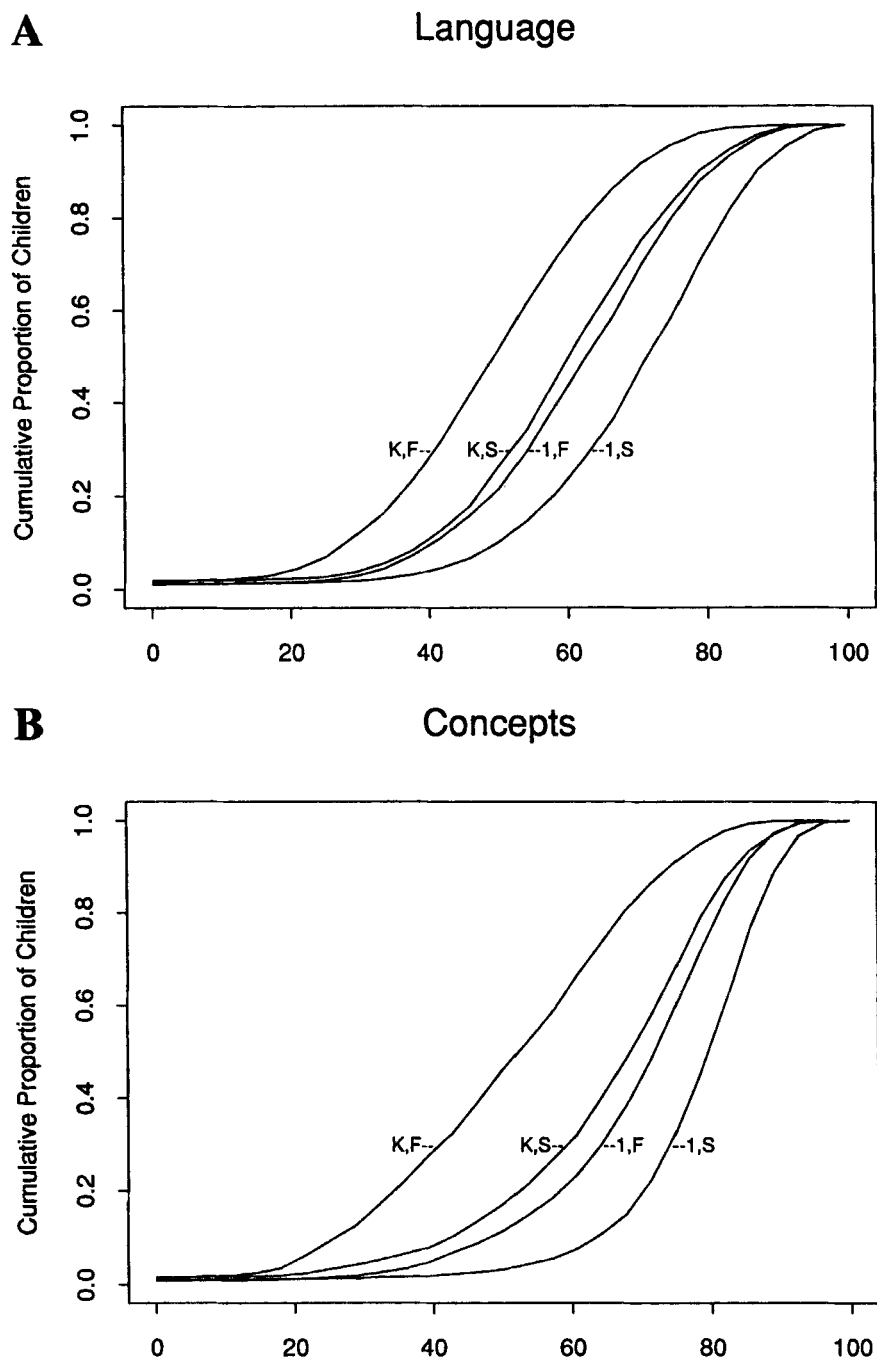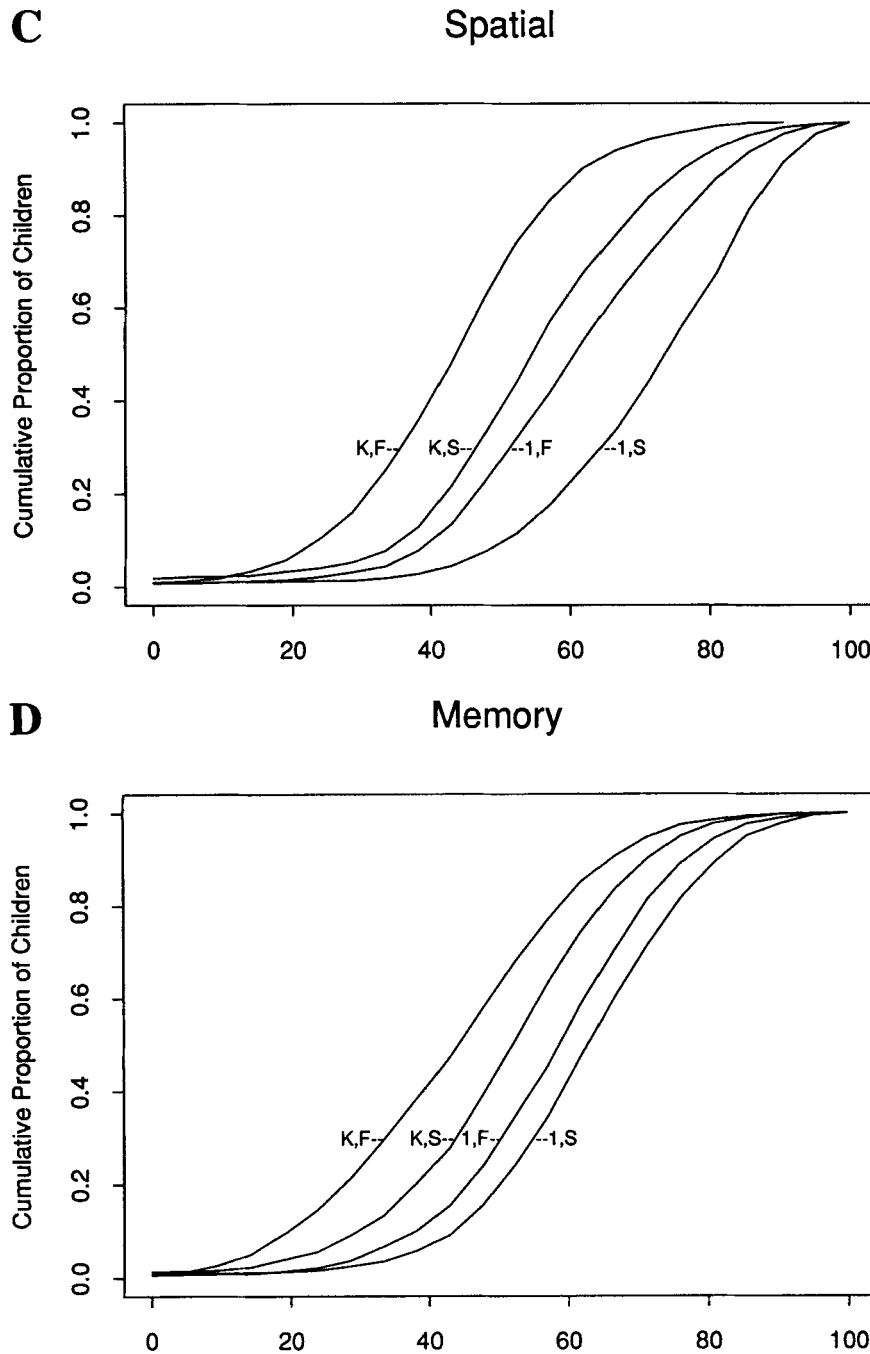## B   Concepts

Figure 2   Cumulative proportion of children receiving scores of at least certain levels for (A) Language, (B) Concepts, (C) Spatial Operations, and (D) Memory (K,F = Kindergarten, Fall; K, S = Kindergarten, Summer; 1, F = Grade 1, Fall; 1, S = Grade 1, Spring).

all demonstrate moderate to large positive effects, indicating that growth is greater during the school year than during the summer. For Memory, on the other hand, the contrast appears to show significantly greater growth for the object location task during the summer than during the school year. However, the contrast must be interpreted in context—there is an

accelerating downward trend, so that the contrast in fact compares moderate summer growth (0.39) with the average of a moderate kindergarten year growth rate (0.36) and a much smaller first-grade growth rate (0.15). The unexpected direction of the significant contrast, then, is an artifact of inappropriate averaging. Moreover, the actual magnitudes of school-

C

Spatial

Cumulative Proportion of Children

K,F-- K,S-- --1,F   --1,S

0   20   40   60   80   100

D

Memory

Cumulative Proportion of Children

K,F-- K,S--1,F-   --1,S

0   20   40   60   80   100

summer differences for both Memory subscales are so small as to be of little theoretical interest; the statistical power associated with the extremely large sample size of the study is such that effects with magnitudes below 0.20 could be found to be significant at the .05 level.

DISCUSSION

In this study, we have found that the growth of certain basic cognitive skills in kindergarten and first-grade children is related to variation in environmental input. We used a time-period comparison design in which growth in a single population was examined over time periods of equal length which varied in input. Our study was cross-sectional, with different individuals tested at each time point. However, the fact that the extent of change in skills, averaged over children, shows an uneven pattern of growth demonstrates input sensitivity over the population studied. A direct evaluation of the causal relation between input and growth in individuals would

Table 3 Differences between School versus Summer Change with Confidence Intervals

	Mean School Year Change minus Summer Change	95% Confidence Interval
Language	.783	.602–.964
Concepts	.772	.600–.944
Spatial operations	.522	.352–.692
Memory	−.092	−.274–.090

Table 5 School Year Growth versus Summer Growth for Subscales

	Contrast	F	Pr > F
Language:			
Vocabulary	.803	73.99	.0001
Syntax	.527	30.36	.0001
Conceptual:			
Categorical	.652	50.13	.0001
Spatial	.717	63.85	.0001
Spatial:			
Sequencing	.504	33.90	.0001
Integration	.377	15.86	.0001
Memory:			
Location-object	−.259	7.48	.0062
Sound-picture	.277	8.22	.0042

Note: Degrees of freedom for all F statistics are (1, 7879).

require a longitudinal study in which the same children were tested at each time point. In such a study it would also be possible to examine summer growth in relation to growth over the school year under conditions where the socioeconomic status and type of summer activities of each individual are known.

We found large differences in growth for periods spent mainly at school versus periods spent mainly at home in three of the four domains studied, Language, Spatial Operations, and Concepts, but not in Associative Memory. Growth of language and of spatial skill is often attributed to intellectual potential, and we briefly discuss our findings of input sensitivity in relation to what is currently known about de-

Table 4 Raw and Standardized Subscale Scores for Each Group

	Language				Concepts			
	Vocabulary		Syntax		Knowledge-Based		Spatial	
	Percent Correct	Mean z Score	Percent Correct	Mean z Score	Percent Correct	Mean z Score	Percent Correct	Mean z Score
Fall K	49	−.578 (.963)	53	−.458 (.927)	49	−.619 (.993)	56	−.714 (1.016)
Spring K	61	.004 (.932)	61	−.056 (.965)	61	−.033 (.955)	72	−.003 (.944)
Fall 1	63	.101 (.914)	64	.076 (.936)	64	.138 (.871)	76	.184 (.859)
Spring 1	72	.514 (.877)	72	.466 (.961)	72	.544 (.814)	85	.564 (.689)

	Spatial Operations				Memory			
	Sequencing		Integration		Location-Object		Sound-Picture	
	Percent Correct	Mean z Score	Percent Correct	Mean z Score	Percent Correct	Mean z Score	Percent Correct	Mean z Score
Fall K	32	−.692 (.671)	58	−.538 (1.076)	42	−.507 (1.015)	53	−.422 (.990)
Spring K	46	−.181 (.899)	67	−.038 (.983)	50	−.150 (.931)	61	−.078 (.980)
Fall 1	55	.137 (.945)	70	−.038 (.859)	58	.238 (.909)	65	.106 (.928)
Spring 1	72	.764 (.871)	75	.431 (.810)	61	.392 (.883)	72	.407 (.926)

Note: Quantities in parentheses are standard deviations.

velopment in these domains. We also briefly discuss the conceptual and associative memory domains.

Input and Language Growth

It is commonly believed that language skills reflect, in large part, the intellectual potential of language learners. The arguments are different for vocabulary and syntax. For vocabulary, there are large individual differences (cf. McCarthy, 1954). The centrality of vocabulary subtests on IQ tests (e.g., Wechsler, 1989) reflects the assumption that what is critical to skill levels is variation in the ability to learn from language input rather than variation in input itself. Vocabulary is assessed in achievement tests as well as IQ tests, indicating a recognition that vocabulary is also an acquired skill, reflecting, at the least, a "crystallized" form of intelligence. Earlier studies provide evidence of a substantial relation between amount of language input and vocabulary growth (e.g., Huttenlocher et al., 1991). The present study suggests a causal role of language input in vocabulary growth. So too does Ramey and Campbell's (1984) work showing that children from impoverished backgrounds who were given preschool interventions that emphasized language had higher verbal IQ's than randomly selected controls.

For syntax, it has been widely claimed that there are not substantial individual differences in skill levels. Syntactic development is viewed as largely controlled by species-wide biological endowment (e.g., Lenneberg, 1967; Pinker, 1984). Although some input is posited as necessary to "trigger" such an innate mechanism, it is believed there should be little variation in growth across a wide range of input. Perhaps because of this belief, there has been little systematic exploration either of the extent of individual differences in syntactic skill or of the possible relation of such differences to variations in input. The existing empirical work showing input effects on syntactic growth focuses on the emergence of particular syntactic forms (e.g., grammatical morphemes or auxiliaries) in relation to the frequency of those forms by parents (e.g., Brown, 1973; Newport, Gleitman, & Gleitman, 1977).

The present study shows substantial input effects on the growth of syntax as well as vocabulary in kindergarten and first-grade children. Whereas the findings of input sensitivity for vocabulary verify those of earlier studies, the findings for syntax are new. The syntax task used in the present study involved the comprehension of various syntactic forms. Earlier studies of syntactic development, focusing on commonalities in development, have examined production, and in younger children. However, the results of the present study are supported by evidence from a study of production in somewhat younger children. Huttenlocher (1997) found large individual differences in the complexity of the utterances produced by different 4-year-olds and a substantial relation to the complexity of syntactic input by their caregivers.

Input and Spatial Growth

Systematic evidence that naturally occurring variations in input affect the growth of spatial skill in children is lacking in the existing literature. Large individual and population differences in skill levels for spatial operations exist, notably between the sexes on spatial transformation tasks (cf. Maccoby & Jacklin, 1974). It is commonly believed that individual and group differences in skill levels, including sex differences, have a biological basis (e.g., Beatty, 1979). It also is known that there are differences in environmental input to boys and girls that could be relevant to spatial growth (e.g., Astin, 1974). The only evidence in the existing literature showing that input can affect skill levels is based on intervention studies with adults. Whereas these studies show that training does indeed increase levels of spatial skill, the training given has been on tasks that are very similar to the tasks used to assess spatial skill levels (cf. Baenninger & Newcombe, 1989).

The present study suggests that naturally occurring variations in input are causally related to spatial skill levels. Not only does the finding of substantial school effects provide convincing evidence that spatial growth is input sensitive, it also suggests what the critical input may consist of, such as activities that are more likely to occur at school than at home (such as geometric instruction, activity with puzzles, and so on).

For the Spatial Scale, the proportion of growth at home versus school varied with start level. Children in the top quartile gained no more at school than at home, although children in the other three quartiles gained more at school. Possibly, children with higher start levels participate in more school-like spatial activities at home than other children, so that the input at school may be redundant. Alternatively, children with greater potential might learn relatively more from home activities than other children.

Input and Conceptual Growth

For the concepts scale, we did not start out to assess the effects of environmental input. Rather, in de-

veloping scales to assess language skills and spatial skills, we wanted to determine if object and spatial concepts were critical. We found that the concepts tasks did not form scales with other language or spatial tasks, but did form a single scale with one another.

The object concepts task involved grouping objects based on their functions or underlying characteristics, thus capturing conceptual knowledge such as that assessed in both ability and achievement tests using tasks that are verbal in nature. Conceptual knowledge, whether it is assessed verbally or nonverbally, might be expected to reflect the extent of input children receive. Whereas input at school would seem relevant to the growth of object concepts, this is not so clear for spatial concepts. Such concepts could, at least in principle, arise directly from observation of the spatial relations among objects without input from caregivers. The fact that school has substantial effects on growth for spatial concepts suggests that conventional input is important for their acquisition as well as for the acquisition of more knowledge-based object concepts.

The Concepts Scale, like the Language Scale, showed very large differences in growth for school versus home. Further, differences in growth between school and home were similar across starting levels. Indeed, the same aspects of input may be involved in both of these domains of cognitive growth. In particular, the critical input to language and conceptual growth may be the curriculum at school, which provides to children a major source of new information as well as the forms to express that information linguistically. Here, as for language, the effects of input at home, although not as great by this age as the effects of school input, may have been greater at an earlier age.

Lack of Relation of Input to Associative Memory Growth

For the Associative Memory Scale, as for the Concepts Scale, we did not start out to assess effects of environmental input. Rather, the original purpose of the two associative memory tasks (sound-object associations and location-object associations) was to examine whether these skills were critical to language and spatial skills. We found that they were not; instead, the two associative memory tasks were highly related to one another.

Consistent with the existing literature (cf. Kail, 1979), our findings show that memory develops over the 5 to 7 year age period. The literature provides evidence that there are strategy changes in certain memory tasks over this age period; notably, increases

in rehearsal in short-term memory tasks (e.g., Flavell, Friedrichs, & Hoyt, 1970). It seems likely that such age-related changes might be differentially sensitive to school input. Indeed, Morrison et al. (1995) and Ramey and Campbell (1984) both show input effects on memory span for the names of a set of pictures; these tasks are similar to those for which rehearsal changes have been demonstrated. Our memory task, where no school effects appear, differs from these tasks in that it involves learning new paired associates as opposed to requiring the use of words learned earlier. It should be noted that the time-period comparison design reveals only input effects that are differentially greater during periods at school than periods at home. Thus the overall age-related increases we find may be environmentally driven, but the environment at home may be as important as that at school. More work is clearly needed to examine how memory mechanisms may be affected by various sorts of input.

CONCLUSIONS

The present study demonstrates a substantial relation between naturally occurring variations in input and growth in the spatial and language domains, as well as in the conceptual domain. We assessed four large groups of children drawn from a single population. Thus, although the children tested at each time point differed, the large data set was used to infer growth in the population. Direct assessment of input sensitivity in individuals would involve following those individuals over the entire time period. This would allow examination of the relation of the particular types of experiences individuals receive over different time periods to the growth of individual cognitive skills over those periods.

Clearly, there are great commonalities in cognitive growth across children, reflecting the underlying propensities of human beings to construct certain skills from universally available aspects of the environment. Regardless of the importance of these common aspects of development, it is useful to focus on the variability among individuals, because the cognitive skills important for successful functioning are not exhibited equally across people. Such individual differences may reflect the particular biological characteristics of individuals as they develop within varying environments. Those differences that are sensitive to variations in the input individuals receive are of special interest in that children's environments can be altered, with potential consequences for both individuals and social groups. The present study shows a substantial relation between input variations and growth in types of skill where such a relation had not

previously been demonstrated in an unambiguous way.

ADDRESSES AND AFFILIATIONS

Corresponding author: Janellen Huttenlocher, University of Chicago, Department of Psychology, 5848 South University Avenue, Chicago, IL 60637; e-mail: hutt@cicero.uchicago.edu. Susan Levine is also at the University of Chicago; and Jack Vevea is at the University of North Carolina, Chapel Hill.

APPENDIX

In developing the PTCS, we began with eight tasks, four of which we believed might potentially tap a language factor and four of which we believed might potentially tap a spatial factor. We designed each task so that, insofar as possible, it would require only the skill being tested, without requiring extraneous skills which might vary, thus obscuring the child's ability on the target skill. Analysis of the results for the eight tasks revealed four distinct domains. Here we describe the tasks we began with, as they were all included in the final scale.

The four language tasks included syntax, vocabulary, object concepts, and verbal association tasks:

For the *syntax* items, a sentence was read aloud, and children were required to choose among a set of pictures. The pictures showed the same elements interacting in different ways. The sentences included a range of forms—passives, negatives, complex sentences.

For the *vocabulary* items, a word was read aloud and children chose among contrasting pictures (the picture of the object named). The pictures showed closely related meanings. All words were object names varying in frequency in the language. (Object names were used to avoid relational words that are closely tied to syntactic processes because of the possibility of producing a correlation between skills that might be distinct.)

For the *object concepts* items, children were shown a picture of a target object. They matched it to one of a set of four objects, all perceptually different from the target, on the basis of being an instance of the same concept. (Language skill requires semantic/conceptual distinctions, and children's nonverbal concepts might be highly related to language skill.)

For the *verbal memory* items, a set of nonsense syllables was said aloud while nonsense objects were presented (e.g., this is a vek, this is a zop). Children then chose the proper picture from a set of four, after being given one of the nonsense syllables (e.g., which is the vek?). Difficulty was manipulated by varying the number of pairs presented prior to the probe nonsense word. (Language skill requires acquisition of large numbers of words, and the ease of acquiring such pairings might be highly related to scores on vocabulary or even syntax tests.)

The four spatial tasks included spatial transformation, spatial sequencing, spatial concepts, and spatial association tasks:

For the *spatial transformation* items, children were shown fragments of a geometric shape resulting from dividing the shape along major axes of symmetry. They chose from among four shapes the one that could be made by translating and/or rotating the fragments.

For *spatial sequencing*, children were shown a repeating pattern of simple geometric forms. They chose from among four alternatives, each indicating a possible next shape or set of shapes in the sequence.

For *spatial concepts* items, children were shown a picture of two objects in a particular spatial relation to one another. They matched that picture to one of four choices of which only one depicted the same spatial relation. (Relating objects to one another spatially is necessary to perform spatial operations, and skill levels might be highly related to scores on spatial operations.)

For the *spatial association* items, children were shown a large circular array of familiar objects in particular locations around the perimeter. Then a circle was presented with blank spaces at each of the locations. A particular location was pointed out, and children indicated the object that had appeared at that location from a set of four objects. (Spatial operations require memory for the locations of items, and memory skill might be highly related to scores on spatial operations.)

The eight tasks were analyzed using methods of item response theory and Q3 analyses (Yen, 1984) to examine the dimensionality of groups of items and establish which of the tasks formed unidimensional scales. Vocabulary and syntax tasks proved to be unidimensional, forming a language scale. Spatial integration and sequencing tasks also were unidimensional, forming a scale of spatial operations. The category and spatial concepts tasks did not form unidimensional scales with the language or spatial tasks, respectively, but did unite to form a single concepts scale. The verbal memory and spatial memory tasks did not form unidimensional scales with the language or spatial tasks, but did satisfactorily combine with each other to form an associative memory scale. Correlations among the resulting four scales were: concepts and associative memory, .50; concepts and language, .61; concepts and spatial operations, .60; associative memory and language, .42; associative memory and spatial operations, .49; and language and spatial operations, .52 (all $p < .001$).

REFERENCES

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Reading comprehension and education* (pp. 77–117). Newark, DE: International Reading Association.

Astin, H. S. (1974). Sex differences in mathematical and scientific precocity. In J. C. Stanley, D. P. Keating, & L. H. Fox (Eds.), *Mathematical talent: Discovery, description, and development*. Baltimore: Johns Hopkins University Press.

Baenninger, M., & Newcombe, N. (1989). The role of experi-

ence on spatial test performance: A meta-analysis. *Sex Roles, 20,* 327–344.

Baltes, P. B., & Reinert, G. (1969). Cohort effects in cognitive development of children as revealed by cross-sectional sequences. *Developmental Psychology, 1,* 169–177.

Beatty, W. W. (1979). Gonadal hormones and sex differences in nonreproductive behaviors in rodents: Organizational and activational influences. *Hormones and Behavior, 12,* 112–163.

Bisanz, J., Morrison, F. J., & Dunn, M. (1995). Effects of age and schooling on the acquisition of elementary quantitative skills. *Developmental Psychology, 31,* 221–236.

Bouchard, T., Lykken, D. T., McGue, M., Segal, N. L., et al. (1990). Sources of human psychological differences: The Minnesota study of twins reared apart. *Science, 250,* 223–228.

Brown, R. W. (1973). *A first language: The early stages.* Cambridge, MA: Harvard University Press.

Campbell, F. A., & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. *Child Development, 65,* 684–698.

Campbell, F. A., & Ramey, C. T. (1995). Cognitive and school outcomes for high-risk African-American students at middle adolescence: Positive effects of early intervention. *American Educational Research Journal, 32,* 743–772.

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Psychological Bulletin, 27,* 702–722.

Clarke-Stewart, K. A. (1973). Interactions between mothers and their young children: Characteristics and consequences (pp. 1–109). *Monographs of the Society for Research in Child Development, 38.*

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research, 66,* 227–268.

Diamond, M. C., Krech, E. L., & Rosenzweig, M. R. (1964). The effects of an enriched environment on the biology of the rat cerebral cortex. *Journal of Comparative Neurology, 123,* 111–119.

Entwistle, D. R., & Alexander, K. L. (1992). Summer setback: Race, poverty, school composition and mathematics achievement in the first two years of school. *American Sociological Review, 52,* 72–84.

Flavell, J. H., Friedrichs, A. G., & Hoyt, J. D. (1970). Developmental changes in memorization processes. *Cognitive Psychology, 1,* 324–340.

Gardner, H. 1983. *Frames of mind: The theory of multiple intelligences.* New York: Basic Books.

Gottfried, A. W. (Ed.). (1984). *Home environment and early cognitive development.* New York: Academic Press.

Gottfried, A. W., Gottfried, A. E., Bathurst, K., & Guerin, D. W. (1994). *Gifted IQ: Early developmental aspects* (The Fullerton Longitudinal Study). New York: Plenum Press.

Greenough, W. T., & Volkmar, F. R. (1973). Pattern of dendritic branching in occipital lobes of rats reared in complex environments. *Experimental Neurology, 40,* 491–504.

Guilford, J. P. (1967). *The nature of human intelligence.* New York: McGraw-Hill.

Heyns, B. (1978). *Summer learning and the effects of schooling.* New York: Academic Press.

Honig, S. A. (1982). Parent involvement in early childhood education. In B. Spodek (Ed.), *Handbook of research in early childhood education.* New York: Free Press.

Huttenlocher, J. (1997). Environmental effects on cognitive growth. Paper presented at the biennial meeting of the Society for Research in Child Development, Washington, DC.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27,* 236–248.

Huttenlocher, J., & Levine, S. C. (1990). *Primary Test of Cognitive Skills.* Monterey, CA: MacMillan-McGraw-Hill, California Test Bureau.

Kail, R. (1979). *The development of memory in children.* New York: W. H. Freeman.

Kamin, L. (1974). *The science and politics of IQ.* Hillsdale, NJ: Erlbaum.

Karmiloff-Smith, A. (1992). *Beyond modularity.* Cambridge, MA: MIT Press.

Kellaghan, T., Sloane, K., Alvarez, B., & Bloom, B. S. (1993). *The home environment and school learning.* San Francisco: Jossey-Bass.

Lenneberg, E. (1967). *Biological foundations of language.* New York: Wiley.

Maccoby, E., & Jacklin, C. N. (1974). *The psychology of sex differences.* Stanford, CA: Stanford University Press.

McCarthy, D. (1954). Language development in children. In L. Carmichael (Ed.), *Manual of child development* (2d ed., pp. 492–630). New York: Wiley.

Morrison, F. J., Smith, L. S., & Dow-Ehrensberger, M. (1995). Education and cognitive development: A natural experiment. *Developmental Psychology, 31,* 789–799.

Newport, E. L., Gleitman, H., & Gleitman, L. R. (1977). Mother I'd rather do it myself: Some effects and noneffects of maternal speech style. In C. Snow & C. Ferguson (Eds.), *Talking to children: Language, input and acquisition* (pp. 109–149). Cambridge: Cambridge University Press.

Orshansky, M. (1988a). Counting the poor: Another look at the poverty profile. *Social Security Bulletin, 51,* 25–51.

Orshansky, M. (1988b). Commentary: The poverty measure. *Social Security Bulletin, 51,* 22–24.

Pinker, S. (1984). *Language learnability and language development.* Cambridge, MA: Harvard University Press.

Plomin, R., & DeFries, J. C. (1980). Genetics and intelligence: Recent data. *Intelligence, 4,* 15–24.

Ramey, C. T., & Campbell, F. A. (1984). Preventive education for high-risk children: Cognitive consequences of the Carolina Abecedarian Project. *American Journal of Mental Deficiency, 88,* 515–523.

Scarr, S. (1992). Developmental theories for the 1990s: Development and individual differences. *Child Development, 63,* 1–19.

Snow, C. E., & Ferguson, C. A. (Eds.). (1977). *Talking to children.* Cambridge: Cambridge University Press.

Sperry, R. W., Gazzaniga, M. S., & Bogen, J. E. (1969). Interhemispheric relationships: A preliminary study. *Brain and Language, 37,* 42–58.

Stevenson, H. W., Lee, S. Y., & Stigler, J. W. (1986). Mathematics achievement of Chinese, Japanese, and American children. *Science, 231,* 693–699.

Stigler, J. W., Lee, S. Y., & Stevenson, H. W. (1987). Mathematics classrooms in Japan, Taiwan, and the United States. *Child Development, 58,* 1272–1285.

Teuber, H. L. (1955). Physiological psychology. *Annual Review of Psychology, 6,* 267–296.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *The Stanford-Binet Intelligence Scale: Fourth Edition.* Chicago: Riverside Publishing.

Thurstone, L. L. (1938). *Primary mental abilities* (Psychometric Monographs 1). Chicago: University of Chicago Press.

Wechsler, D. (1989). *Manual, Wechsler preschool and primary scale of intelligence—Revised (WPPSI-R).* San Antonio, TX: Psychological Corp.

Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement, 8,* 125–145.