Education and Cognitive Development: The Evidence from Experimental Research
Author(s): Donald Sharp, Michael Cole, Charles Lave, Herbert P. Ginsburg, Ann L. Brown and Lucia A. French
Source: *Monographs of the Society for Research in Child Development*, Vol. 44, No. 1/2, Education and Cognitive Development: The Evidence from Experimental Research (1979), pp. 1-112
Published by: Wiley on behalf of the Society for Research in Child Development
Stable URL: http://www.jstor.org/stable/3181586
Accessed: 28-05-2017 23:10 UTC

MONOGRAPHS OF THE
SOCIETY FOR RESEARCH IN
CHILD DEVELOPMENT

SERIAL NO. 178, VOL. 44, NOS. 1–2

EDUCATION AND COGNITIVE DEVELOPMENT: THE EVIDENCE FROM EXPERIMENTAL RESEARCH

DONALD SHARP
TICUL, YUCATAN

MICHAEL COLE
UNIVERSITY OF CALIFORNIA, SAN DIEGO

CHARLES LAVE
UNIVERSITY OF CALIFORNIA, IRVINE

WITH COMMENTARY BY
HERBERT P. GINSBURG
ANN L. BROWN AND LUCIA A. FRENCH

AND REPLY BY
MICHAEL COLE

MONOGRAPHS OF THE SOCIETY FOR RESEARCH IN CHILD
DEVELOPMENT, SERIAL NO. 178, NOS. 1–2

CONTENTS

ABSTRACT

SHARP, DONALD; COLE, MICHAEL; and LAVE, CHARLES. Education and Cognitive Development: The Evidence from Experimental Research. With Commentary by HERBERT P. GINSBURG, and by ANN L. BROWN AND LUCIA A. FRENCH. *Monographs of the Society for Research in Child Development*, 1979, 44(1–2, Serial No. 178)

This *Monograph* reports the results of a series of experimental studies and a sociodemographic survey designed to determine the relative influence of age and educational experience in the development of cognitive skills as manifested in formal, psychological experiments.

The research was conducted in rural Yucatan, Mexico, primarily with Indio (Maya) and Mestizo subjects ranging in age from approximately 10 to 56 years. The experimental tasks centered on (1) the categorization of objects and words using a variety of procedures, (2) memory for objects and words, and (3) problem solving using verbal and nonverbal materials. The survey, which included 393 informants, investigated a number of demographic and economic factors associated with the amount of schooling obtained by people in the population which was the target of the experimental studies. An adapted IQ test subscale was also included in the survey questionnaire.

Results of the experimental studies strongly indicated that age and educational experience contributed differentially to performance on different tasks. For many of the tasks studied, performance increased as a function of education but not with age. For other tasks, both increasing age and educational experience were associated with increased performance, while for a few tasks performance improved as a function of age but was not affected by differences in education. Education exerts its strongest influence on tasks where taxonomic principles are the criterion of correct classification, where correct performance requires the subject to provide structure to the nominal stimulus array, where problems must be treated hypothetically without recourse to real-world knowledge, or where problem contents are clearly school related (e.g., arithmetic calculation). Insofar as functional criteria and real-world knowledge can produce a correct response or the task is well

structured, education-related effects are reduced, and age-related effects are more prominent.

Multiple-regression analysis of the sociodemographic survey showed that availability of schools in the informant's hometown was the major constraint on the amount of education obtained. Residual variation in schooling obtained was explained largely by economic factors. The IQ predicted a small but significant amount of the residual variation in schooling consistent in magnitude with other studies. The force of these survey results was to rule out massive selection artifacts as sufficient explanation of the performance differences associated with education.

The final section of the *Monograph* discusses the implications of these results for developmental-psychological theory and the general problem of assessing the cognitive impact of education using standard psychological testing procedures. It is pointed out that conclusions regarding age-related differences in several kinds of cognitive performance from studies conducted in the United States and other highly developed countries should be re-examined for the extent to which their results arise from age-correlated differences in educational experience. Difficulties in specifying the generality of the cognitive changes under study is related to the difficulty of finding experimental techniques which retain analytic power *and* ecological validity.

I. INTRODUCTION

For both practical and theoretical reasons, the last decade has seen a renewed academic discussion of the intellectual consequences of formal education. On an international scale, training in basic academic skills was promoted as an essential prerequisite to economic progress (Gray 1956); nationally, the expenditures for Operation Head Start were often justified as a means of breaking the "cycle of poverty" by teaching young children the basic intellectual skills that promote learning to read and do arithmetic; following such influential theorists as Hunt (1961), these latter programs assumed that intellectual skills could be taught and that schools were the place where they were most likely to be learned, at least for the poor and "culturally different."

More recently, the assumptions that formal education promotes economic development and that intellectual skills are learned in school have come under attack.

At both the national (see Greer 1972) and international levels (Harmon 1974), evidence has mounted that the development of formal educational facilities follows rather than leads economic development.

In a parallel manner, many social scientists have come to accept the idea that schools do not modify children; rather, they screen children according to previous (if not predetermined) levels of ability. This is by no means a new idea. Many years ago Pillsbury (1920) actually advocated large-scale intelligence testing precisely because tests could carry out the school's screening function efficiently. Jencks's (1972) widely cited conclusions concerning the inability of schools to change educational outcomes predicted from home background factors, when combined with popular interpretations of Jensen's (1971) conclusion that "educability" is inherited, have also contributed substantially to the notion that schools do *not* bring about basic cognitive changes.

These issues are a matter of very broad concern, and the controversy surrounding them is easy to understand. Unfortunately, they are also very

The research reported here was supported by grants from the Office of Education and the Carnegie Corporation.

difficult issues to resolve, scientifically as well as in terms of social policy.

One major impediment to their scientific resolution is the close correlations that exist in modern society between contending causal factors associated with test performance, scholastic achievement, and the economic consequences of formal education. In the United States, age and years of education are almost perfectly correlated over the ages of 5–15 years, since almost all children start school about age 5 and continue for 10 or more years. Socioeconomic status and years of education are also closely linked, as are socioeconomic status and race, both of which correlate highly with educational attainment and standardized test performance. A multitude of factors interact to determine the later economic effects of education to the individual and society.

Cross-cultural research offers one possible avenue of approach to the problems of interpretation inherent in domestic research by addressing the same issues in circumstances where the correlations among socioeconomic conditions, age, and education are either absent or different from those observed in domestic research on this topic.

Unhappily, cross-cultural research introduces its own impediments to resolving disputes about the effects of experience (including educational experience) on the development of cognitive abilities, even in cases where a strong case can be made for interpreting the significance of the contributions of independent variables.

These difficulties center on the interpretation of behavior in experimental tests where the subjects have, by virtue of the experiential differences which led to the investigation in the first place, differential familiarity with the aims and contents of the test. As we have been at pains to point out on other occasions, there may be barriers to assuring the equal representativeness of task elements in cross-cultural research that cannot be overcome within the confines of standard, experiment-based research (see Cole 1975; Cole & Scribner 1974; Scribner 1976a). The effect of these difficulties is to narrow the range of inferences that we will want to draw from cross-cultural data, an issue to which we will return at the end of this *Monograph*.

PURPOSE OF PRESENT RESEARCH

This *Monograph* follows that tradition of cross-cultural research which uses other cultural settings as an opportunity to de-confound theoretically important independent variables. In this case we will concentrate on formal education as a causal factor in mental development.

Research by Greenfield carried out more than a decade ago (see Greenfield & Bruner 1966) pointed to schooling as a prerequisite to the development of various logical operations and classificatory skills. Greenfield took advantage of the unequal distribution of educational facilities in Senegal to disentangle age and educational experience. Since that time, many have followed Greenfield's basic strategy, and substantial evidence has accrued in support of the generalization that schooling teaches a variety of the skills which produce improved performance on psychological tests of intellectual ability (see for reviews Cole & Scribner 1974; Dasen 1977; Scribner & Cole 1973).

While there is widespread evidence that formal educational experience does indeed inculcate the skills needed to perform well on psychological tests (our proxy for "cognitive development"), the data are by no means unanimous. As the reviews cited above make clear, for some tasks and some cultural groups formal schooling makes little difference.

Our own experience (Cole, Gay, Glick, & Sharp 1971) and our reading of the literature suggested several problems which need attention. Paramount at the time we undertook this research were the following:

Range of Age and Educational Experience

While it has been possible to disentangle age and education in cross-cultural studies, the very recent advent of formal education in many of the societies studied has resulted in limited free variation in age and education. For example, in our research in Liberia, we were unable to find tribal adults with any significant educational experience, and the number of available subjects with secondary school experience was drastically limited. This same difficulty permeates the literature.

Social Correlates of Amount of Education

Where schools have been located without respect to internal social and economic divisions (as was the case in parts of rural Liberia in the 1960s), we have some justification for assuming more or less random assignment of children to the major contrast groups—schooled and nonschooled—at least at the lower-grade levels. However, when we turn to places where mass schooling has been available for some time, or when we become concerned with higher levels of education, we clearly want to evaluate exactly those social factors which are known to influence educational achievement in the United States (e.g., family incomes, home language) as well as factors specific to the locale of the research but which may not have exact equivalents here (traditional occupations, availability of schooling).

In short, we must combine the currently accepted practice of comparing educated and noneducated subjects with a study of selective factors which shape their inclusion in these groups.

GENERAL PLAN AND RESEARCH SETTING

With these goals in mind, we undertook a two-part study of education and its cognitive consequences on the Yucatan peninsula in Mexico. The first part of the research consisted of more than a dozen series of studies of various psychological tasks which previous cross-cultural research had shown to vary with educational experience. The second part of the study consisted of a demographic survey designed to specify the selective factors operating to determine the amount of education obtained by Yucatecans. Note that in concentrating our attention in this manner, we are not pursuing a cross-cultural research strategy which draws its experimental manipulations from indigenous cultural practice as advocated in Cole and Scribner (1975) or Scribner (1976a). The consequences of the present research strategy will be considered in Chapter VI.

Our choice of Yucatan for a study aimed at the effects of schooling was guided by several factors, paramount among which was the great variability in educational attainment of its heterogeneous population combined with historical circumstances which made at least moderate levels of schooling available to a significant proportion of its rural, adult population.

The Area

The Yucatan peninsula lies in the extreme southeast of Mexico and juts north separating the Gulf of Mexico from the Caribbean Sea. This region of Mexico, which includes the states of Yucatan, Campechi, and Quintana Roo, prides itself on its ancient cultural heritage (where the concept of zero, developed trade routes, written language, and great archeological complexes were a few of the features present) and its historical independence of Mexico City. Yucatec Mayan (thought by many to be the oldest of the seven presently spoken Mayan languages) is still the dominant language in much of the countryside. It is the only language, or the first and preferred language, of a significant proportion of the rural adult inhabitants of the peninsula. Active resentment of the federal government was still not uncommon at the time our research was conducted, although the fratricidal "Caste War," which began in 1847, was officially put down almost 100 years ago.

Although some light industry and a tourist trade exist in Merida, the capitol, the newly constructed tourist complexes of Quintana Roo, and a few of the larger towns, 50%–60% of the peninsula's inhabitants earn their livelihood directly or indirectly from agriculture. The major food crops are corn, beans, and citrus fruits, while the major cash crops are henikin (sisal) in the northern and western areas in the State of Yucatan, and chicle and precious woods (most notably mahogany and tropical cedar) in the State of Quintana Roo. Towns vary in size from more than 200,000 (Merida) to tiny "ranchos" of only a few houses run by a single family or a small group of related families.

Virtually all of our research was conducted in 37 cities and towns extending from Merida in the northwest of the peninsula along Mexico Route 184 to Chetumal in the extreme southeast. Figure 1 presents a sum-

Characteristics

	Towns Rank Ordered by Rated Traditionality	Higher education grades > 9	Secondary schools grades 7-9	Small industry	Population > 5000	Medical facilities	Central market	Telephone service 1971	Agencies; Businesses	Telegraph service	Post office 1971	Restaurants	Population < 80% Maya	Potable water 1971	Paved road 1971	Home crafts for sale	Colonial churches	Bus service 1971	Migrant work by people	Electricity 1971-72	Stores, general	1000 ≤ population ≤ 5000	Primary school thru 6th	Population > 80% Maya	Pure agriculture focus	Maya predominantly spoken	Primary < grade 6	Nonliteracy > 60%	Population < 1000	Number of interviews in demographic survey
1.	Merida	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X									
2.	Chetumal	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X									
3.	Bacalar	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X									80
4.	Ticul	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X									154
5.	Tekax	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X									
6.	Peto		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X									
7.	Muna		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X									14
8.	K-50		X				X	X	X		X	X	X	X	X		X	X	X	X	X	X								
9.	Teabo				X	X		X			X	X	X	X	X	X	X	X	X	X	X		X							14
10.	Sacalun												X	X			X	X	X	X	X	X	X		X					13
11.	Mani											X	X		X	X	X	X	X	X	X	X	X							12
12.	Pustunich													X	X	X	X	X	X	X	X	X	X	X						13
13.	Chapab																X	X	X	X	X	X	X	X	X					18
14.	Yotholin												X		X	X	X	X	X	X	X	X	X							11
15.	Dziuche								X				X	X	X			X	X	X	X	X	X	X						
16.	Dzan													X		X	X	X	X	X	X	X	X	X						17
17.	K-71											X	X	X	X			X	X	X	X	X	X		X					
18.	Catmis												X	X	X	X		X	X	X	X	X	X		X					
19.	Mayapan															X	X		X	X	X	X	X	X	X	X		X		17
20.	Nohbek														X					X	X	X	X	X	X					
21.	Cocoa																			X	X			X	X		X		X	12
22.	San Jose														X	X	X	X	X	X	X			X	X	X	X	X		15
23.	Chaksikin									X							X		X	X	X	X	X	X	X					
24.	Xyatil																			X	X	X	X	X	X					
25.	Yunku																	X	X	X	X			X	X		X	X	X	21
26.	San Felipe																			X	X			X	X	X	X	X	X	
27.	San Antonio																	X	X	X	X			X	X	X	X	X	X	21
28.	Senor																			X	X	X	X	X	X	X				
29.	X-hazil																			X	X			X	X	X	X	X	X	
30.	Kancabchen																X		X	X	X	X	X	X	X	X				
31.	Xoy																			X	X			X	X	X	X	X	X	
32.	Ramonal																				X			X	X	X	X	X	X	
33.	Reforma																							X	X	X	X	X	X	
34.	Xpichil																						X	X	X	X		X	X	
35.	Tuzik																							X	X	X	X	X	X	
36.	Dos Aquadas																							X	X	X	X	X	X	
37.	Cinco de Mayo																							X	X	X	X	X	X	

FIGURE 1.—Town characteristics as a function of towns' rated traditionality

mary of several features distinguishing these towns which are also marked on the accompanying map (fig. 2). The order in which the towns are listed in the table reflects their traditionality based on 15 separate rank orderings collected from a single research assistant who had worked extensively in each of the towns. The rank orderings were collected about once a month for a period of a year. The use of a single informant in collecting data of this type presents problems of which the authors are well aware but which could not be avoided in this case since only one of our assistants was familiar with all of the towns used for our studies. This limitation notwithstanding, there

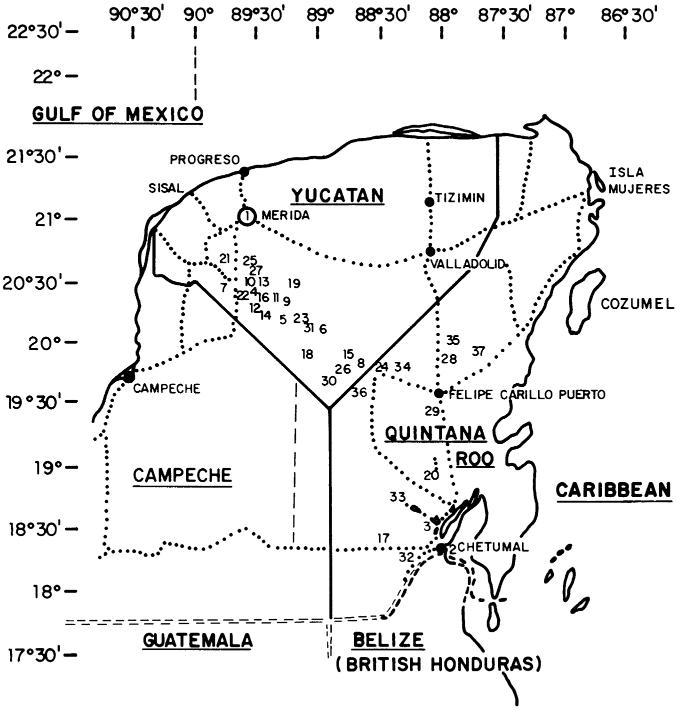

FIGURE 2.—Map of the area with towns noted. The towns are numbered corresponding to order in fig. 1.

is a reason to believe that the composite rank ordering (consisting of an ordering based on the mean rank order for each of the 37 towns on the 15 separate rank orderings) is highly reliable. Test-retest Spearman correlations for all 15 rank orderings ranged from +.79 to +.98, with a mean of +.89. Furthermore, Spearman correlations between this informant's ratings and ratings from three other assistants for towns known to all four informants were always greater than .80.

The features listed in figure 1 were selected in the following manner. Working with an informant other than the one who had done the rank orderings, a list of more than 50 characteristics thought to reflect town sophistication was constructed. During the course of the work on the demographic survey (to be reported in detail later), we visited each of the towns and rated them for the presence or absence of a given characteristic (using local informants other than our assistants). In the case of contradictions between informants, the group consensus was taken as the value of presence or absence of that characteristic. Exceedingly general or specific items were eliminated from analysis, and a resulting list of 28 characteristics retained. It is this list that appears in figure 1. As a final step, the characteristics were rearranged according to the similarity of the characteristics present and absent (see Stefflre, Reich, & McClaran 1971) to give the composite figure presented in figure 1, where the pattern appears relatively clear-cut. Two major groupings appear: the large (population > 5,000) towns and cities with some light industry, central markets, and complete social services (features marked in the upper left-hand corner of the table); and the small, predominantly agricultural, Mayan communities (marked in the lower right of the fig.). The towns in between may be considered transitional and are in many cases satellites of the first group of towns. For example, Sacalum, Mani, Pustunich, Chapab, and Yotholin were all satellites of Ticul in the sense that government administration is centered in Ticul and local residents depend on Ticul for necessities not available in their home towns. These variations in sizes, ethnic composition, and occupation of the inhabitants coincide with a number of other variations among inhabitants of the peninsula which we will explore in our educational-demographic survey.

The Schools

As figure 1 makes clear, one of the key features that reflects the sophistication of a town is the level of education available to its inhabitants (the correlation between grade of school in town and our scale of town modernity was +.88 for our sample of towns). Naturally, aside from being an indicator of the sophistication of a town, schools are also an environment in and of themselves for those that attend them. Physically, the schools may show a great deal of variability. The typical rural school may be a three-room building with more than one grade taught in a room, while the typical urban school may closely approximate standards in the United States. But in terms of structure of the educational system and content of the curriculum, everything is remarkably constant throughout the Republic.[1]

The following curriculum may be considered representative of the structure of the educational system at the time our research was conducted (1970–1974).

Primary school curriculum (grades 1–6) is uniform throughout the Republic in that no matter what the institution—federal, state, private, or church schools—every student receives the same basic curriculum, which may be divided into the traditional categories: (1) reading (in Spanish); (2) mathematics; (3) the study of nature; (4) geography; (5) civics, geography, and history.[2] Much of the first year is devoted to reading, but interlaced with

[1] Note that in what follows emphasis is on uniformity and sophistication of content; this does not imply uniform sophistication of presentation. In many cases rote presentation of material for memorization is the rule, especially among the older generation of teachers. In the discussion that follows the general outline of the governmental educational program is emphasized. There are many subprograms tailored to specific regional needs and numerous private academies offering a host of specialized courses of study.

[2] Note that, as this report is being written, the educational system is undergoing a massive revision, both structurally and in presentation of content. Our description pertains to the system in effect at the time of our studies.

the readings are materials relevant to the major areas of study that will emerge as separate entities in the later grades. Basic concepts such as fractions and areas in mathematics, or public health in the study of nature, are introduced to the student in the very earliest grades. With each succeeding year the basic concepts are repeated and new concepts added so that, by grade 6, the student is supposed to have the basic framework that will allow him to continue his education or will serve him in good stead for day-to-day practical life in Mexico. Under this system, a student who has passed through grade 3 has been exposed to such concepts in mathematics as: larger, smaller, equal to; tens and hundreds units, thousands; addition, subtraction, multiplication and division of three-digit numbers with remainders; fractions; Mexican monetary system; metric system; measures of area and volume; geometric figures; and solid figures.

By grade 6 all this information has been integrated into a general base where the student, through his mathematics, has progressed to simple algebraic formulations; in the study of nature he has a rudimentary knowledge of public health and ecology and a highly detailed knowledge of the structure of his own body and the complete biological classification system. In reading he has progressed from four-line poems by Garcia Lorca in grade 1 to topical essays in both Spanish-speaking and foreign authors in grade 6. Similar examples could be given for the remaining areas of study.

Upon completion of primary school, the student, if he is able to, passes on to secondary school (grades 7–9). Grade 7 is generally an intensive review of the primary curriculum and may be considered an extension of the first 6 years. Even in primary school, repetition from year to year is the rule, especially in science and mathematics. This repetition is important because many students cannot regularly attend school and would be hopelessly lost if considerable overlap of material were not provided. In grades 8 and 9 the student must choose between two programs of study—a vocational (technological) or a preparatory (academic) course of study—the first giving the student a skill (carpenter, electrician, dress designer, e.g.) and the second providing the base for further study. The programs are not mutually exclusive, and the system is flexible enough to allow for transfer from one program to the other. In the preparatory course, the student, by the time he has finished grade 9, has had trigonometry, algebra, plane and solid geometry, world history, social sciences, and national language, and may have elected one or more languages and a selection of more specialized electives.

Upon the completion of grade 9 the student may continue to one of three advanced schools or enroll in a host of specialized (generally private) academies. The major alternatives in the government program are: (1) *preparatoria* (a course of 3 years consisting of grades 10–12 to prepare for professional [university] studies), (2) *technologia* (a course of 3–4 years in specialized technical schools, producing technicians in a variety of areas including medicine, agriculture, geology, as well as in advanced work for the grades taught in secondary school), and (3) *normal* (a course of 3–4 years producing lower-grade level elementary school teachers). All three areas are open to secondary students who have elected the *preparatoria* program, while only certain of the *technologica* programs are open to the graduates without further study on the secondary level. With completion of the 3-year preparatory program, the student progresses directly to professional studies in the university. A liberal arts program as we know it in the United States is generally not available in the national system but may be available from private academies. Five of the towns and cities that were sites of our investigations had one or more of these higher educational institutions. Merida has all of the above including university studies on the professional level; Chetumal has a *technologica* and *preparatoria*; Ticul has a *preparatoria* course, Tekax has an agricultural *technologica*, and Bacalar, a *normal* for education of primary school teachers. In addition, each of these towns, along with Peto, Muna, and Kilometro-50, has one or more secondary schools. For the remainder of our towns, education was limited to six full grades or less.

Some Additional Background Factors

As we will show in more detail later, while the younger generation of Yucatecans is better educated than older generations, the distribution of educational experience both within as well as between generations is quite marked. This heterogeneity, as can be seen from the general description of our intentions, was a basic resource in our work. In addition to differential availability of education, the educational situation is complicated by the fact that in many areas a language problem exists with the children coming to school speaking only Maya, while their teachers speak only Spanish.

Beyond the mere fact of school availability and language knowledge, economic factors appear to be paramount in determining how long a child attends school. During the period in which this work was conducted, the minimum daily wage was roughly 28 pesos ($2.24 American), giving the average worker in Yucatan a monthly income of roughly $58 a month. The equivalent figure for Quintana Roo was roughly $86. These figures are somewhat misleading because they assume that a person is employed (generally by the government or a corporation) and works a 5½-day workweek, assumptions that are not justified for a large number of our subjects or subjects' parents. Many people were farmers or artisans working on a piecework basis who received no fixed wages; others worked only a limited number of days a week, or a few hours a day. Furthermore, minimum salary figures generally refer to concessions and businesses under direct or indirect government control and are in large part inapplicable to the many small businesses that constitute the vast majority in the area under consideration. Farmers' incomes vary greatly, both as individuals, and as groups, but the average is certainly not as high as the "average" wage earner.

The unnumbered table below represents an estimate of the cost per day

Grade	Expense/Day (Pesos)	Total Cost/Month Pesos	Total Cost/Month Dollars
1, 2	1.00 personal 2.00 supplies	60.00	= $ 4.80
3, 4, 5	2.00 personal 3.00 supplies	100.00	= $ 8.00
6	2.00 personal 5.00 supplies 1.15 3 uniforms	162.50	= $13.00
7	2.00 personal 5.00 supplies 1.15 3 uniforms	162.50	= $13.00
8	8.00 personal 8.00 supplies 1.15 3 uniforms	342.50	= $27.36
9	8.00 personal 10.00 supplies 1.15 3 uniforms	382.50	= $30.50

and month for one child in each of the grades 1–9 based on interviews conducted in 1973. These figures represent a well-equipped student. A child can always get along with two uniforms instead of three, nothing to eat, and other people's books. Making allowances, it is still clear that the cost of education was substantial at the time of our research relative to income, increasing dramatically after the second grade, with a major jump coming in the sixth grade and during secondary school.

One of the major sources of increase in secondary school is the cost of books. Textbooks are provided gratis for all elementary school students by the government but must be paid for in secondary school. Although there are a limited number of scholarships available in secondary schools, substantial government aid in the form of scholarships is not available until after the ninth grade. Thus, for workers of even relatively high income, secondary school becomes a real watershed, especially if the student comes from a large family (average family size of those interviewed in the educational survey reported below was seven, i.e., two adults and five children) where many children may be studying simultaneously. For those families of lower income, such as farmers and campesinos, every year of successive education beyond grades 2 and 3 may be an overwhelming burden. These economic factors, when combined with problems of language and the availability of schools, combine to make school attendance quite variable among the Yucatecan population. Although virtually everyone between 7 and 25 years of age from larger towns has attended school at some time, many older people in all areas and many younger adults from rural areas have never done so. Moreover, as we will show in some detail in a later section, many people, either by virtue of economic necessity or the absence of educational facilities, have attended school for a very short time—sometimes fewer years than is necessary to become completely literate in Spanish.

With this brief description of the social and economic milieu of our research, we will turn to the experiments on cognitive functioning.

II. EXPERIMENTAL STUDIES: STUDIES OF CATEGORIZATION

The experiments fall into three broad classes. The first group investigated the structure of people's lexical knowledge as that knowledge manifested itself in a variety of categorization tasks.

There were two central reasons to begin by investigating natural language categories. Although the focus of our research was the consequences of education, we knew we were dealing with a bilingual population in which the "mix" of Spanish and Maya varied greatly. Educational experience (all in Spanish) was one of the factors influencing the use of Maya and Spanish, so we needed materials where the two languages exhibited lexical organization of the same structure, or we could confuse the consequences of language knowledge and education.

We also knew that we would need to be sure of the status of lexical knowledge where categorization per se was the dominant feature of the task. We planned to use these data in studies which assume a given category structure and go on to determine the way that category knowledge is used in more complex cognitive tasks.

Following this logic, the second series of studies concentrated on memory processes. Our earlier research (Cole et al. 1971) had suggested far-reaching consequences of school experience for the way people go about remembering discrete sets of items. The present studies extended those observations, concentrating on the search for evidence that subjects were deliberately employing mnemonic strategies, including category clustering, to maximize performance.

The final set of studies investigated verbal and visual problem solving, where categorization and memory are two of the important constituents of performance.

The initial studies were begun in the winter of 1970 with the aid of a trilingual Mayan (English, Spanish, and Yucatec Maya) informant who had previously conducted anthropological and linguistic research in the area. Our basic procedure was to visit a town in the region, stopping to talk with the mayor, federal deputy, or the schoolteacher and explain our business. Operating through responsible village leaders, we obtained permission to interview people and conduct our experiments. Since we were working in rural areas, we often had to concentrate our work with adults in the evening hours or weekends. We saw children in school, or in the courtyard if no spare room was available. Adults were paid for their participation, the amount depending upon the time involved. Children were given candy.

We occasionally met resistance of two kinds in work with adults (children were almost always anxious to participate). In traditional Mayan villages, many adults mistrusted and feared outsiders; in such cases, we took extra care to establish rapport, but even so some people refused to cooperate with us. We also found it more difficult to obtain the cooperation of traditional, generally older, adult women among whom formal conversation with strangers is frowned upon. These difficulties induced us to give up on our attempt to include equal numbers of males and females in our studies where males predominate. We did evaluate sex differences within groups, and significant performance differences will be reported where they occur.

Another problem in establishing basic experimental groups arises from an intrinsic aspect of the heterogeneous population that we had chosen to study. The local dialects (both Maya and Spanish) contain three terms descriptive of locally recognized subcultures: "Indio," "Mestizo," and "Ladino." Each term is relatively easy to characterize as a stereotype. "Indio" refers to uneducated, monolingual Maya speakers who may have some familiarity with Spanish but maintain Maya as their preferred language in virtually all social interaction. Most people in this group work the land (campesinos), either as paid labor on large estates or work cooperatively owned land (ejidatarios).

"Mestizo" generally refers to people of Mayan heritage who have had some education and speak Spanish at least well enough for normal communicative purposes; they are likely to be involved directly to varying degrees in the money economy. "Ladinos" are well educated, and while they may speak some Maya they are likely to be descendants of early colonial settlers, citizens from other parts of the republic, or people of Mestizo background who have achieved relatively high social status, have come to de-emphasize their Indian heritage and identify with mainstream Mexican national culture. This category does not appear often in our studies, but it must be realized that many of our student subjects, particularly at the higher educational levels, would fall into this group.

The borders of these stereotypes are not clear. An "Indio" (referred to hereafter as "Maya" in designating groups) may have experienced a little schooling, many "Mestizos" have both Spanish and Maya names, and there are prominent "Ladinos" with Mayan surnames and rural backgrounds.

A good deal of pilot work suggested that education is a key determinant of how people label themselves and others, but clearly occupation, income, and language are a part of the process, as are the situations in which the terms are used.

In the studies to be described, we usually adopted three population labels—Maya, Mestizo, and Student—the first two of which correspond to the first two labels just discussed, and the latter of which is used in its conventional meaning—someone currently enrolled in school. These subdivisions allowed us to concentrate on differences in educational experience but left us vulnerable to the problems of covarying factors (language, occupation, etc.) which are obvious from the defining characteristics of the groups.

We attempted to deal with this problem in two ways. First, we conducted experiments in which one or more of the potentially influential covarying background factors was evaluated experimentally. Thus, for example, we conducted some of our work in Maya (for half the subjects) and Spanish for the remainder. Second, when the resources became available, we conducted our educational-demographic survey to get a better idea of how various background factors interact in the general population under study. We were then able, in the case of a few of our experimental studies, to include key survey questions as independent variables in the research design. Unhappily, our ignorance combined with constraints on time and personnel led us to begin the experimental work prior to completion of the survey. Luckily, the results of the survey, and those studies conducted using survey questions, implicated education, and not covarying population characteristics, as the key to performance differences.

CLASSIFICATION OF COMMON OBJECTS

All natural languages possess lexicons in which individual words are recognized as instances of the same type (Berlin, Breedlove, & Raven 1973); Yucatec Maya (McClaran 1971) and Spanish are no exception to this rule. We know, however, that the organization of some lexical domains differs significantly between Yucatec Maya and Spanish in ways that systematically affect cognitive performance (e.g., recognition memory in the color domain: Stefflre, Vales, & Morely 1966). Consequently, we chose to study domains where both class memberships and their organization were extremely similar for the two languages.

A number of investigators (see Cole & Scribner [1974] or Greenfield [1966, pp. 270–318] for summaries) have found that uneducated people are less likely than their educated counterparts to partition groups of objects from different semantic classes into static taxonomic sets, although the object labels are familiar and their lexical organization is known to the subject (e.g., he will give these objects as instances of the category). Instead, objects are grouped according to functional relations or visual similarity. Operating with

Polaroid photographs of objects commonly encountered in rural Yucatan, we sought in our first study to determine the influence of education on the tendency to reproduce the potential semantic relations in a free classification task.

Subjects.[3]—Twenty subjects were drawn from each of the following four population groups, except the adults of whom there were 47: (1) 9–12-year-old students (grade = 3, age = 11 years); (2) 12–15-year-old students (grade = 6, age = 13 years); (3) 13–18-year-old secondary school students (grade = 7.8, age = 16 years); (4) adults, 47 monolingual Maya speakers (grade = 0, age = 38 years).

Stimulus materials and procedures.—Color Polaroid photographs of the following 20 common objects (determined by pilot work and questioning our informants) were spread on a table in front of the subject: animals (bull, dog, duck, horse, and turkey), utensils (pot, plate, bottle, cup, knife), food (tortilla, coconut, corn, onion, bread), and clothing (shirt, hat, pants, sandals, and hrupil—a native dress).

Each subject was read the following set of instructions:

> I have here pictures of things that are very well known to you and to all of us [here the experimenter pointed to the photographs that were spread out in front of the subject]. What I want you to do is first name the things that appear on the cards for me, then I want you to take the cards and put them in groups that you feel go together, or are alike, or are the same in some way. Do you understand what I want?

Subjects experienced little difficulty in working with the photographs (a problem encountered in some previous research) or in understanding what was required of them. The few subjects who showed evidence of difficulty were read the instructions again and allowed to ask questions for clarification. No one showed any difficulty after a second reading of the instructions. The experiment was conducted in Spanish for everyone but the monolingual Maya speakers.

Results.—The data were aggregated by group to obtain an adjacency matrix, reflecting the number of times each item was placed in the same group as each other item.

The raw frequencies with which items appear together are a very cumbersome means of summarizing the results. Fortunately, a variety of multidimensional scaling techniques exist for analyzing data of this type. Our results will be presented in two ways. Figure 3 contains a graphic plot of the degree of semantically defined group cohesion. The plot is a combination of outputs from the Johnson hierarchical clustering program (an algorithm that produces a tree diagram of items in a similarity matrix starting from the most specific and proceeding in steps to the most general

[3] Numbers given in grade and age here and in later studies are group averages.

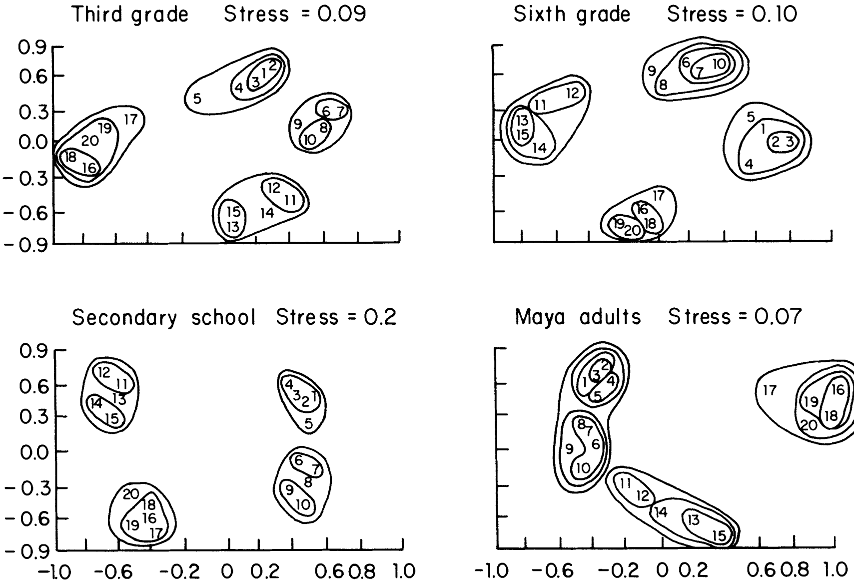

FIGURE 3.—Two-dimensional representations of classifications produced for 20 objects used in sorting and memory studies. The numbers correspond to the following objects: 1 = pot, 2 = plate, 3 = cup, 4 = knife, 5 = bottle, 6 = bread, 7 = tortilla, 8 = cebolla (onion), 9 = coconut, 10 = corn, 11 = turkey, 12 = duck, 13 = dog, 14 = bull, 15 = horse, 16 = shirt, 17 = huipil, 18 = trousers, 19 = hat, 20 = shoes (sandals).

[Johnson 1967]), and the Torsca nonmetric multidimensional scaling program (Young & Torgerson 1967). This program accepts as input a matrix of similarity scores such as we have for our common everyday items and gives an output of Cartesian coordinates for each item. The number of dimensions to be used, which can vary from one to the number of items, is specified by the investigator. In many cases the investigator is not certain as to the number of dimensions that might underlie a given set of items. In such cases he has the option of having the program calculate the interpoint distances first for a large number of dimensions, and then repeat the procedure for a smaller and smaller number until a criterion of acceptability is achieved. The criterion of acceptability is usually reflected in a jump in the value of a statistic termed "stress," reflecting the goodness of fit of the obtained structure for any given number of dimensions to the similarity relations inherent in the original matrix. The larger the value of the stress statistic, which can vary from 0.0 to 1.0, the worse the fit between the interpoint distances in the geometrical structure and the original similarity scores. By comparing the stress scores for various dimensional solutions, the investigator can determine when an increase in "stress" does not merit further simplification of the data. As figure 3 indicates, the stress values for all four groups were below the generally accepted "good" value of .30 for this number of items.

This two-dimensional solution also permitted us to combine the Johnson and "Torsca" representations demonstrating complementary representations of the group categories.

Inspection of the separate solutions for the two adult groups, both in terms of configuration, indicated that the groups were virtually identical; consequently, they are aggregated on the graph. The extremely low stress value of .07 for the combined groups (the lowest obtained) supports this decision since, if the groups were radically different, a much higher value would have been expected. The placement of items in four graphs, then, represents the Torsca multidimensional scaling solutions for each of the population groups. The items are indicated by numbers whose identity is contained in the figure legend. The concentric rings that surround the numbers grouping them into clusters represent nodes from the output of the Johnson hierarchical clustering program—the smaller rings representing more specific relations, while larger and larger rings indicate increasingly more general relations. Although there is rather convincing grouping by category, there are clear differences between groups both in the differentiation within categories and the separation between categories. The between-category differences are apparent, for example, in the way that the adults place animals close to foods, while these two categories are widely separated for the secondary school students.

The concentric rings around sets of items as well as differences in the distance between items reflect within-category organization. Here the tightness of the organization of the schoolchildren's sorting is quite apparent. Note that not only do the Mayan noneducated adults place animals and foods near each other in the scaling analysis, foods and utensils are a part of the same hierarchy (e.g., there is a ring encompassing them).

There are also clear differences in the "tightness" of the categories among the schooled groups. The older (and more educated) children give more tightly organized grouping.

The source of differences in category cohesion is more clearly illustrated by an examination of table 1. In this table we have recorded the total number of times that subjects within each subject population placed items from a given category with items from that same or other categories. Inspection of the four panels in the table indicates clearly that the relative proportion of within-category sorting increases as a function of grade: within-category sorting of the uneducated adults is at a level equal to that of the third graders.

Examination of particular entries in the various panels of table 1 sheds more light on the categorizing principles at work. For the category "clothing," virtually all sorting is *within-category* for all the population groups. But for the remaining categories there is a tendency to combine members of one category with members of another in ways that strongly suggest a functional

TABLE 1

RESULTS OF PICTURE-SORTING STUDY: *N* ITEMS SORTED TOGETHER WITH ITEMS OF THE SAME OR OTHER CLASSES

	Animals	Food	Clothing	Utensils
		Grade 3 Students[a]		
Animals	150	66	0	13
Food	...	157	6	109
Clothing	...	...	170	26
Utensils	...	...	...	151
		Grade 6 Students[b]		
Animals	141	13	0	0
Food	...	135	7	39
Clothing	...	...	190	8
Utensils	...	...	...	144
		Secondary Students[c]		
Animals	149	10	16	10
Food	...	146	23	51
Clothing	...	...	154	18
Utensils	...	...	...	147
		Adults[a]		
Animals	275	238	42	81
Food	...	230	13	413
Clothing	...	...	392	31
Utensils	...	...	...	330

[a] Ratio of within-category response to total responses = .61.
[b] Ratio of within-category response to total responses = .90.
[c] Ratio of within-category response to total responses = .83.

basis for categorizing. This functional sorting decreases with increasing educational experience. Thus, for example, the noneducated adults are more likely to place a food item and a utensil together than they are to put two food items or two utensils together in a particular grouping. Mixing of food items and animals is also common. Further analyses showed that this mixing occurred because some of the animals (duck, bull, and turkey) can plausibly be responded to as food.

While there is a slight tendency to observe the same "errors" among secondary school students and sixth graders, there is relatively little variation among categories for these groups.

Conclusions.—All of the groups included in this study demonstrated clear classification along taxonomic lines. This result is important as a demonstration that the category membership of the object label whose category membership has been established by linguistic criteria will be manifested in the context of a psychological task that directly solicits categorization.

Although these observations were collected as "baseline" data on a sample of populations to be included in later studies (e.g., populations varying markedly in age and education) and hence are not entirely appropriate for "age versus education" comparisons, the data do suggest that increased educational experience increases the application of taxonomic category information to the explicit categorization of objects. This conclusion would be strengthened if we had included adult groups which served as matches for the educational level of the schoolchildren as we did in later studies. Although Mayan language may have influenced those subjects who responded in Maya, there are no linguistic data or data from our later studies to support such a conclusion.

CATEGORIZING AND RECATEGORIZING WITH MULTIPLE CRITERIA

In the previous study, where categorizing per se was the task, a small set of category names could fully partition the alternatives. In fact, use of the category names was the only way in which a single set of criteria could exhaustively group the objects. The next two studies were directed at instances where several partitions of the set were possible using adjectival category labels (red, large) instead of nouns which subsume nameable instances of natural categories. The stimuli in the first of these studies were pictures of geometric shapes (triangles and squares). We were interested not only in which attributes would be selected as a criterion for sorting the stimuli but whether or not subjects would sort on more than one dimension if asked to do so on successive probes.

A number of investigators have found that educated subjects are more likely to use form as a dimension for sorting geometric stimuli of this kind. In addition, they are more likely to be able to re-sort the set once they have attained an initial classification (see Cole & Scribner [1974] and Greenfield & Bruner [1966] for summaries and discussions of these data). No studies have included adults with varying amounts of education in addition to children of school age, a strategy that avoids the confounding involved in the previous sorting study using pictures.

Two studies will be described. In the first, all of the stimuli are sets of eight cards. The sets of cards differ only in the dimensions and values of dimensions of the particular stimuli. In the second study, responses to cards depicting abstract forms are compared with responses in a task involving maize (a culturally relevant stimulus; see, e.g., Irwin and McLaughlin [1970]).

Sorting Geometric Stimuli

Subjects.—Thirty-two subjects, roughly half of whom were male, represented each of the groups included in the present study: (1) first-grade students (7–9 years of age), (2) third-grade students (9–10 years of age), (3) sixth-grade students (12–13 years of age), (4) Mestizo bilingual adults with 3 years or less of formal education (grade = 1.5, age = 18 years).

Stimulus materials.—Two sets of stimuli were made of figures colored on the blank side of 3 × 5 index cards with a nylon-tip marker pen. For the first set, the cards varied in color (red vs. yellow), number (1 vs. 2), and form (triangle vs. square). For the second set, the color and number dimensions remained the same, but size (large vs. small circles) replaced form as a variable dimension.

Procedures.—Subjects were seated opposite the experimenter at a low table upon which one of the sets of cards had been placed. They were then given the following instructions, a variation of those used in the picture-sorting study described previously:

> I have here a group of cards with figures on them. [Here the first group of cards was spread out so that the subject could see the figures on the face of the cards.] What I want you to do is take the cards and put them into groups [or piles] that you feel go together or are alike, or are the same in some way. Do you understand what I want?

If the subjects experienced difficulty or indicated that they did not understand, the instructions were repeated. If the subject still did not understand after a second presentation of the instructions, he was dropped from the experiment. It was necessary to exclude only three subjects, two from the first grade and one from the third.

When the subject had attempted to sort each of the card sets, one of his initial groupings was reconstructed on the table, and he was told:

> This is the way that you put these cards into groups the first time. What I want you to do now is show me if there is another way to put the cards in groups that you feel "go together" or are alike or are the same in some way.

The subject's re-sort was then recorded, and the procedure repeated for the second set of cards. Counterbalancing was observed throughout, so that half of the subjects received the form set first, the other half the size set.

Results.—Data for the first sorts are presented in table 2 in terms of the number of subjects (out of 32 possible in each group) who sorted correctly on at least one of the dimensions. Although the most common correct grouping was for a subject to sort two piles of four cards each along one of the dimensions present in the task, some subjects further subdivided along a second dimension (i.e., groups of red and yellow, e.g., were further subdivided into red-large/red-small vs. yellow-large/yellow-small). In such a case, subjects were scored as sorting both color and size (e.g., as sorting on two dimensions). As a result of this possibility, "Total Dimensions" could exceed

TABLE 2

N Correct Classifications, First Choice

A. Form Cards

	Color	*N*	Form	Total Correct Sorts	Total Dimensions
Grade 1	3	1	4	6	8
Grade 3	8	3	7	16	18
Grade 6	10	6	17	25	33
Mestizo adults	4	5	9	14	18

B. Size Cards

	Color	*N*	Size	Total Correct Sorts	Total Dimensions
Grade 1	3	2	0	5	5
Grade 3	6	7	4	15	17
Grade 6	15	12	8	29	35
Mestizo adults	3	6	5	11	14

"Total Correct Sorts" (which only allowed for one dimension per trial and no further subdivisions). The column labeled "Total Correct Sorts" indicates the number of subjects in each group who were able to group the cards according to at least one dimension on the first try.

The data indicate that the ability to classify either set of these cards increases primarily as a function of grade in school rather than with age per se. There is a clear increase in the total number of consistent sorts over grades, $\chi^2(3) = 11.98$ and 20.3 for the form and size cards, respectively, $p < .01$, but the Mestizo adults sort at a level slightly but not significantly lower than the performance of the third graders, $\chi^2 = .13$ and .62 for the two sets, respectively, although they are the oldest subjects in the study. Their performance is slightly but not significantly superior to the first graders, $\chi^2(1) = 3.2$ and 2.25, $p > .10$.

The data do not support any conclusions regarding a color-to-form, $\chi^2(3) = 4.0$, $p > .10$, or color-to-number shift in the dominant dimension selected for sorting as a function of either age or education (χ^2 for the color-number shift on the two sets equals 1.2 and 0.0); the two youngest and two oldest were combined in this test to avoid excessively small cell frequencies.

The data for subjects' reclassification is contained in table 3. A sixth column has been added to the table to record the conditional probability of a reclassification given a correct first classification. Reclassifications of stimuli closely parallel the data from table 2. The conditional probabilities demonstrate that reclassification also increases with years of formal education: the Mestizo adults reclassify at a level approximating that of the third graders but, in this case, more clearly superior to the first graders.

TABLE 3

RECLASSIFICATION OF STIMULI

A. FORM CARDS

	Color	N	Form	Total Sorts	Total Dimensions	P(r)/S[a]
Grade 1	6	0	1	7	7	.14
Grade 3	5	4	4	13	13	.69
Grade 6	9	2	8	19	19	.89
Mestizo adults	3	0	9	12	12	.58

B. SIZE CARDS

	Color	N	Size	Total Sorts	Total Dimensions	P(r)/S
Grade 1	5	1	2	8	8	.00
Grade 3	7	4	5	15	16	.36
Grade 6	7	3	9	19	19	.95
Mestizo adults	4	4	4	12	12	.55

[a] Probability that cards are resorted on a new dimension given a first, consistent sort.

Sorting Geometric Figures and Maize

This study, a partial replication of the previous study, was conducted to evaluate two additional questions. First, a group of monolingual Mayan adults with no educational experience was added to the groups represented in the original study. Second, eight pieces of Yucatec maize were substituted for the size cards used in the previous study. Yucatec maize kernels are red or yellow in color and vary rather markedly in size, yielding two dimensions comparable to the cards. In addition, some kernels come off the cob as "twins" or "doubles" according to the Yucatec expression. With the exception that the two sets of eight items consisted of cards and corn, instead of two sets of abstract figures on cards, the materials and procedures were exactly the same as the previous study. The choice of stimulus objects was prompted by studies (e.g., Gay & Cole 1967; Irwin & McLaughlin 1970) demonstrating that the use of "culturally appropriate" or familiar stimulus objects could enhance performance, or wipe out performance differences between groups where they existed. Since there is also negative evidence on this question (Greenfield 1974), the problem seemed to warrant further attention.

Subjects.—The groups consisted of 32 subjects each, roughly half of whom were male, chosen from the following populations: (1) first-grade students (7–8 years old); (2) third-grade students (9–11 years old); (3) sixth-grade students (12–13 years old); (4) Mestizo adults with some primary education (grade = 3.0, age = 20.7 years); (5) noneducated, Mayan-speaking adults (grade = 0.0, age = 37.8 years).

Procedure.—The instructions, procedures, and scoring methods remained the same in this study as in the previous one, except that Mestizo adults were permitted to work in either Spanish or Maya, while the Maya adults worked exclusively in Maya.

Results.—The data for the first sorts are shown in table 4: table 4A contains the data for the color/number/form of cards, table 4B the data for the color/number/size of corn kernels.

The data from table 4A represent a rather neat replication of the comparable conditions in the previous study. Here we confirm the importance of education. The difference between grades is significant for the form card stimulus set, $\chi^2(2) = 7.43, p < .05$. It is in the same direction but not significant for the maize, $\chi^2(2) = 4.23, p > .10$. The Mestizo adults perform like the third graders, and the Mayan adults sort least of all. This is the pattern we would expect if education rather than age were a critical factor in performance on this task.

Examination of table 4A quickly tells us that using the corn kernels, the "culturally appropriate" material, does not markedly affect these subjects' ability to sort consistently. The total number of subjects in each of the population groups who consistently sort on the first try is very similar for the two types of stimuli. Moreover, the number who sort the two dimensions (color and number) common to the two studies is also very similar. As we noted in the first study, there is a tendency for subjects presented geometric stimuli to partition groups of four into two subgroups. This subdividing tendency increases with grade for the cards (reflected in the total

TABLE 4

Correct First-Choice Classifications

A. Form Cards

	Color	N	Form	Total Correct Sorts	Total Dimensions
Grade 1	5	3	3	8	11
Grade 3	8	5	2	12	15
Grade 6	7	7	18	22	32
Mestizo adults	4	8	11	16	22
Mayan adults	1	1	1	3	3

B. Maize Kernels

	Color	N	Size	Total Correct Sorts	Total Dimensions
Grade 1	5	1	3	9	9
Grade 3	8	3	3	14	14
Grade 6	10	11	3	20	24
Mestizo adults	7	7	4	14	18
Mayan adults	4	1	1	6	6

dimension score), but for the corn kernels no substantial subdividing is apparent.

The Mestizo adults' data were again analyzed to assess the relation between number of years of education and sorting within the adult group. The result was a reliable partial correlation of grade and performance (with age held constant) of .62(30), $p < .01$. The partial correlation between age and sorting, with the grade held constant, was not significant ($r = .08$).

Conclusion.—These studies reconfirm the influence of years of educational experience on people's ability to impose a classification scheme on stimuli with perceptually conflicting bases of classification.

Our failure to find enhanced performance with indigenous materials replicates the findings of Greenfield (1974), while failing to replicate those of Irwin and McLaughlin (1970). Clearly, familiarity with the contents of the problem is not sufficient to explain all instances of performance deficits on such problems by all cultural groups; rather, something about the arbitrariness of the classification criteria seems to be learned in school—at least in the Yucatan.

CATEGORIZING WITH CONSTRAINT ITEMS SUPPLIED

In the studies just reported the categorizing performance of the basic population groups differed primarily as a function of years of formal education, when it differed at all.

Using pictures of real-world objects with clear taxonomic referents, uneducated adults appeared quite competent at sorting, if somewhat less "taxonomy bound" than schoolchildren. Using artificially constructed stimulus sets with clearly conflicting bases of solution, the less-educated adults experienced difficulties, even with materials for which the classification dimensions were a part of their everyday classificatory repertoire.

What we sought in the present studies were procedures which contained clearly conflicting bases for classification but used real-world stimuli (or their verbal equivalents) belonging to readily identifiable taxonomic classes. In this way, we hoped to learn why we obtained relatively little difference between groups in the first study and substantial differences in the latter two studies.

Category Matching

Birch and Bortner (1966) published a paper directly aimed at the question of categorizing with and without "conflicting" bases for classification of various real-world objects. Their procedure represents one way of combining the two types of studies reported above; they used common objects, and they could systematically vary the conflicting or nonconflicting nature of the choices presented their subjects.

Thus, if the subject was asked to select a shirt, a pot cover, or a palm fiber bag to match with a palm fiber hat, he could choose on the basis of common category or function (shirt), common physical shape (pot cover), or common material (palm fiber bag). If the shirt was included with thread and a small red plate, the potential bases of selection were reduced. The Birch and Bortner procedure fitted our needs so nicely that we undertook a replication of their study using materials appropriate to the Yucatan.

Stimuli.—Table 5 contains a list of the items used in this category-matching study. The items in series A are designed to maximize the conflicting bases according to which an item from the triplet in the right column could be matched with the single item on the left. Those in series B were selected to minimize such conflict.

Subjects.—The subjects were drawn from the five following population groups: (1) first-grade students, 42 in series A (age = 7.8) and 42 in series B (age = 8.0); (2) second-grade students, 38 in series A (age = 8.8) and 32 in series B (age = 8.7); (3) third-grade students, 26 in series A (age = 9.4) and 33 in series B (age = 9.5); (4) sixth-grade students, 20 in series A (age = 13.0) and 20 in series B (age = 13.0); (5) noneducated Mayan adults, 20 in series A (age = 44.1) and 20 in series B (age = 46.6).

Procedures.—Each subject was presented either the A series or the B series. The three "to-be-matched" items were arrayed on one side of a wooden table in front of the subject. The "criterion" item was placed on the other side of the table. Subjects were given the following instructions:

> ———— [subject's name]. I am going to show you some objects that are very well known to you and to all of us. This game requires that you tell me which of these objects ought to be with this other that I will show you. I want also that you tell me, in whatever sense that you know, why you choose as you do. Here you have three objects that are ———, ———, and ———. With which of these three objects should this other go together?

Results.—The results are summarized in table 6 which shows the mean number of categorical matches (out of a possible score of 4) for each of the subject groups and each series along with the percentage of choices which are justified by a conceptual explanation. (The term "conceptual" is borrowed from Birch and Bortner who used a combined criterion which includes both functional and categorical answers.)

In terms of both dependent variables, the proportion of "correct" responses increases significantly as a function of age. The noneducated adults perform very much like the sixth graders who are closest to them in age and furthest in amount of educational experience. The Mayan adults are far superior to the first graders who are closest to them in educational experience. The overall difference between groups is confirmed by analysis of variance, $F = 3.87(4{,}285)$, $p < .01$. The strong visual impression that the A (conflicting) series is more difficult than the B series was confirmed by the analysis of variance, $F = 77.1(1{,}285)$, $p < .01$.[4]

Further evidence of the role of age in determining performance was obtained by a within-group analysis of the first-grade students who varied rather widely in age owing to migration and economic factors which delayed the beginning of schooling for many pupils. For the A series, a significant, positive correlation of .51 was obtained between age and number correct ($N = 42$, $p < .01$). The analogous correlation for series B of .09 was also positive but not significant (performance in this case was near the ceiling, restricting the range of the scores).

Discussion.—The general trend of these results confirms the trend of findings reported by Birch and Bortner (1966) concerning categorical matching with and without competing bases of classification. Of central interest to

TABLE 5

STIMULI USED IN CATEGORY-MATCHING STUDY

	Series A	Series B
1. Palm fiber hat..........	*Shirt* (striped)	*Shirt* (striped)
	Pot cover	White thread
	Palm fiber bag	Small red plate
2. Small red button........	*White thread*	White candle
	Pepsi bottle top	*White thread*
	Small red plate	White chalk
3. Cigarette..............	White candle	Pepsi bottle top
	Mexican match	*Mexican match*
	White chalk	Palm fiber bag
4. Tortilla................	Yellow button	White candle
	Sour orange	*Sour orange*
	Coffee cup	Match

NOTE.—Choices scored as categorical are italicized.

TABLE 6

CHOICES AND JUSTIFICATIONS IN CATEGORY-MATCHING STUDY

	$\bar{X}$ N CORRECT (out of 4 Possible)		% FUNCTIONAL/ CATEGORICAL JUSTIFICATIONS	
GROUP	Series A	Series B	Series A	Series B
Grade 1............................	1.8	3.0	.43	.72
Grade 2............................	2.0	3.4	.57	.79
Grade 3............................	2.5	3.4	.68	.91
Grade 6............................	2.5	3.4	.74	.95
Mayan adults (no education)...........	2.7	3.5	.80	.99

[4] The analysis was based on an unequal N analysis of variance described by Winer (1962, p. 291).

the present research is the evidence that age, not grade, is the key correlate of increased classificatory skill. When we obtained the results summarized in table 6, we hypothesized that the correct, conceptual responses might arise from different mixes of the functional and taxonomic categorization. Birch and Bortner did not distinguish what we have called functional and categorical choices (e.g., "You take the thread and sew on the button" vs. "The button and thread are both sewing things"). Instead, they combined these response types under the general heading "Conceptual." But when we re-scored the data, incorporating a functional-categorical distinction, we found that taxonomic categorization was virtually absent in all groups.

On the surface these results appear to contradict the findings in our previous studies. The resolution is to be found in the different criteria for categorizing. It should be recalled that in our initial categorizing study using pictures of real objects, some functional categorizing was found in all groups and was particularly strong in the older, uneducated group of Maya and Mestizo adults. This result was obtained even though the stimulus set was chosen to match linguistically coded taxonomic categories; functional relations was a byproduct. In the current study, functional relations was clearly salient and counted as correct. It appears, then, that when the responses scored as correct classifications can be produced on the basis of functional criteria, age, not education, will be the controlling independent variable. This suggestion will receive attention in later studies.

Relevant to this interpretation is a subsidiary analysis of the difficulty of the individual problems in which problems 2 and 3 emerged as clearly easier than items 1 and 4. The functional links between "correct" alternatives in problems 2 and 3 are more salient than in problems 1 and 4 which suffer the added disadvantage that the taxonomic links between items scored correct are not at all obvious.

A Triads Test Varying Complexity of Choice

In our replication of Birch and Bortner, the subject's task was to choose one of three objects to match with a previously specified target object. In the current study, subjects were read three words, two of which were always closely related according to semantic criteria (e.g., cow-horse). The third word was either unrelated to the semantic choices in any obvious fashion (cow-stone-horse), or was related to one (cow-milk-horse) or both (cow-milk-goat) of the semantic choices by an obvious functional relation.

Early cross-cultural work by Luria in 1931–1932 (see Luria 1976) showed a strong proclivity on the part of uneducated central Asian peasants to insist upon functional groupings when both functional and semantic classification were possible in a similar verbal-classification task. We suspected, however, that the degree of functional responding would depend upon the degree to which functional and semantic elements in the choice set were in conflict. Such an outcome would be consistent with the results of the previous studies using objects. We also thought that highly educated subjects might be inclined to make semantic choices regardless of the functional possibilities in the problem, although this was by no means a foregone conclusion given the results of the Birch and Bortner replication.

Stimulus materials.—The stimulus materials used in the present experiment are presented in table 7. The first and third columns under the heading "Exemplars" contain items that are related semantically for that trial; all second-column words are "distractors." Problem type 0 represents problems in which the distractor bears no obvious functional relation to either target word. Problem type 1 contains problems in which the middle term bears a functional relation with only one of the two semantically related words. Problem type 2 is constructed so that there are conflicting functional elements relating to both of the semantic choices. For example, it can readily be seen that there are no obvious functional interrelations for the first problem (water-stone-posol) but that in the second problem a chicken lays eggs while a horse does not, and in the third both beans and corn are grown on Mexican farms.

Subjects and procedures.—Subjects were drawn 25 to a group from the following subject populations: (1) highly educated Mestizo adults, ninth grade or above (grade = 10.4, age = 23.3); (2) Mestizo adults with 0–6 years of education (grade = 3.8, age = 37.0); (3) elementary school students (grade = 4.2, age = 10.5).

Treatment of subjects varied only in the order in which two instructional conditions were administered. Subjects were assigned alternately to the similarity-difference and difference-similarity conditions.

TABLE 7

STIMULUS SETS IN TRIAD CLASSIFICATION STUDY

Exemplars			Problem Type
water	stone	posol[a]	0
horse	eggs	chicken	1
corn	farm	beans	2
rifle	corn	trap	0
machete	sheath	ax	1
corn	farm	pepper	2
machete	water	ax	0
rifle	bullet	trap	1
horse	corn	goat	2
machete	water	hoe	0
corn	horse	pepper[b]	1
water	cup	posol	2
machete	sheath	hoe	1

[a] A local drink made from corn.
[b] Chili pepper.

Subjects received the following set of initial instructions:

> I have here a list of words that I will read to you in groups of three. Of these three words there are two words that appear the same among themselves in their meaning. What are they [here the first triad was inserted]? Why?

All subjects were then read the triads in the order listed in table 7, one at a time, and their choices and justifications were recorded in a notebook provided for the purpose. After the list had been completed, the experimenter said:

> Now I am going to read you the same list again, but this time I want you to tell me which of the three words is different and why.

Subjects experienced no difficulty in understanding what was required of them. Responses were recorded verbatim.

Subjects in the differences-similarities condition received the same basic instructions with the specification of "same" and "different" reversed.

Results.—Subjects' responses were coded as "correct" (e.g., the two semantically related items were selected or rejected), or "incorrect," and their justifications were coded according to the following categories:

1. A *semantic* justification, for example, "They are tools," "They are four-footed animals," "They are weapons," "They are food."

2. A justification noting a *common function* or *use*, for example, "They are used to hunt," "They are to drink," "They serve to cut."

3. *Functional entailments*, for example, "The machete goes in the sheath," "The chicken lays eggs," "The goat gives milk."

4. *Specific shared attribute*, as "They are made of metal," or "They both have edges."

5. Finally, an *irrelevant* justification or instances when the subject gave no answer or could not see any connection between the elements in the problem.

The three parts of table 8 provide three measures of performance; the arc sine transformed percentages of correct responses, semantic justifications of responses, and functional choices on error trials. The columns for each population group represent responses to problems with zero, one, or two functional connections between the semantically related items and the distractor.

Part A of table 8 contains the arc sine transformed proportion of correct choices as a function of group, number of potential functional links introduced by the third item in the triad and the choice procedure (same or different). An analysis of variance conducted on these data revealed a main effect of group, $F = 6.8(2,72)$, $p < .01$, choice procedure, $F(1,72) = 19.4$, $p < .01$, and number of functional links, $F(2,144) = 52.7$, $p < .01$. There

TABLE 8

RESULTS OF TRIAD CLASSIFICATION STUDY

	Students (Grade 4.2)			Mestizo Adults (Grade 3.8)			Mestizo Adults (Grade 10.4)		
	A. Transformed Correct Response Score								
N conflicting elements	0	1	2	0	1	2	0	1	2
Similar	2.7	1.6	2.2	2.7	1.7	2.2	3.0	1.9	2.6
Different	2.2	1.5	1.6	2.3	1.6	1.9	2.6	1.9	2.3
	B. Proportion of Semantic Justification of Correct Response								
N conflicting elements	0	1	2	0	1	2	0	1	2
Similar	.02	.05	.14	.14	.14	.28	.24	.24	.47
Different	.02	.03	.05	.26	.14	.18	.37	.19	.23
	C. Proportion of Functional Entailment Responses on Errors								
N conflicting elements	0	1	2	0	1	2	0	1	2
Similar	.50	.92	.91	.31	.87	.88	.60	.92	.99
Different	.45	.65	.33	.44	.67	.48	.67	.64	.66

was also a significant interaction between the choice procedure and the number of functional links, $F = 4.4(4{,}144)$, $p < .01$. From the table it is clear that choice of the taxonomically related exemplar items is greater for the highly educated Mestizo than the two other groups (which have almost exactly identical mean scores) and that choice of these items is greater for similarity than difference judgments but that this latter result is not obtained when one strong functional link is suggested by the third member of the triad (e.g., problem type 1 items in table 8).

Parts B and C provide data on the justifications offered for choices. In view of the large number of seemingly taxonomic choices in part A (a score of 3.0 represents almost perfect choice of the constraint items), it is interesting to note in part B that the level of semantic justification of responses is rather low but that there is a marked increase as a function of *both* age and education. Nonsemantic justifications of the two constraint items (not shown in the table) overwhelmingly named physical attributes in common for the two constraint items (they both have four legs) or a common function (they both work on the farm). Part C shows that the dominant response pattern on error trials for all groups is functional entailment.

Discussion.—The results of this verbal "matching" experiment confirm an influence of education on the tendency to use taxonomic features of stimuli for purposes of classification but also strengthen the accumulated evidence that (*a*) taxonomic categorization is dependent upon the contrast set that is presented, (*b*) response choices which are consistent with a taxonomic classification may be justified in nontaxonomic terms, (*c*) that taxonomically based responding is by no means dominant across stimulus conditions even for a rather highly educated population, and (*d*) that conclusions about the effects of age and education on performance may differ for different features of performance. The strong proclivity to use functional and attribute-specific justifications of responses in this study using word stimuli is completely consistent with the results of the Birch-Bortner replication despite the variation in stimulus materials. The results of the two studies differ in that, to a limited extent, education increases the amount of taxonomic choices and justification of choices in the present case. This apparent difference is most plausibly attributed to the fact that functional and taxonomic criteria enter differentially into the definition of a "correct" response. In the present study the choice between saying "They are food" and "They are to eat" was used to distinguish semantic and functional justifications; the fact that taxonomic *choices* varied strictly with education while justifications scored as "semantic" showed a change with both age and education underscores the fragility of the verbal distinction in a case such as that cited above.

Taken as a whole, the studies described thus far strongly suggest that the influence of education appears when taxonomic criteria are emphasized, while age is associated with increasingly proficient functional classification.

Word-Pair Similarities

As part of the demographic survey to be described later, we used an adaption of the Wechsler Adult Intelligence Scale (WAIS) verbal similarities subtest. In that work our response measure, following the accepted scoring procedures, was a single score which integrated various verbal justifications. By virtue of its purpose in that part of our research, the adapted WAIS scale was given only to adults.

Because the WAIS procedure bears an interesting relation to the "three-word" verbal classification study just reported, we decided to conduct a developmental study contrasting age and grade using these materials. In the standard WAIS procedure two items are presented for judgment, and the subject must tell how they are alike (e.g., "How are an orange and a plum alike?"). It's a "similarity" task. Performance is scored entirely in terms of verbal justifications.

In the present study, we examined the parallel between the typical WAIS verbal classification and the three-words verbal classification study. By accepting cases in which subjects deny any similarity between items as legitimate, and scoring denials of similarity as "incorrect" responses, we could obtain a rough analogue of our measure of a correct match in the three-words study, while retaining the essential features of the WAIS test which uses verbalization as an index of correctness.

Our concern, based on our previous results, was to see if age or education would have the dominant influence on both judgments of similarity and verbal justification of similarity judgments.

Subjects.—The subjects, 12 to a group (half male, half female), were drawn from the following six subject populations: (1) 8–11-year-old students (grade = 3, age = 9.8), (2) 13–15-year-old students (grade = 6, age = 13.8), (3) secondary school students (grade = 7.5, age = 13.8), (4) non-educated Mayan adults (grade = 0, age = 50.0), (5) Mestizo adults with 1–3 years of formal education (grade = 19, age = 27.4), (6) Mestizo adults with 4–6 years of formal education (grade = 5.3, age = 24.8).

Stimulus materials and procedures.—Each subject was read the following list of word pairs in the order in which they are presented below: (1) pine-apple-banana, (2) corn-squash, (3) machete-hoe, (4) pot-frying pan, (5) horse-goat, (6) watermelon-potato, (7) woman-boy, (8) man-horse, (9) honey-water, (10) fly-tree, (11) gasoline-wood, (12) bicycle-automobile, (13) cedar-mahogany, (14) air-water. After a word pair was presented, the subject was asked to explain how he thought they were similar or "appeared like one another." Responses were recorded verbatim.

Results.—The basic results are shown in table 9. Consider first column 1.

TABLE 9

Responses to Word Pairs from Modified WAIS Test

Group	Correct Response (1)	Justification: Semantic (2)	Functional (3)	Attribute (4)	Average WAIS (5)
Grade 3	.40	.13	.05	.82	5.33
Grade 6	.42	.20	.20	.60	7.00
Mayan adults (no education)	.38	.23	.32	.44	6.33
Mestizo adults (0–3 years education)	.50	.32	.27	.42	8.85
Mestizos adults (4–6 years education)	.53	.50	.24	.26	10.83
Students, secondary	.54	.48	.12	.40	11.25

These data represent the proportion of instances in which subjects acknowledge that the two items were similar in some way. Two points stand out. First, there is little spread between scores as a function of age or education. Second, subjects very often rejected the notion that the two items were, or could be, similar. Only the secondary students and adult subjects who had relatively higher educational experience acknowledged such a possibility as much as half the time. In virtually all cases where a subject denied similarity, he indicated a difference in some concrete attribute of one of the items to justify his response (see cols. 2–4).

Column 5 presents the same data in terms of a WAIS score where subject protocols were scored according to the accepted procedures of awarding 2 points to a categorical response or to a high-level generalization (e.g., "They are both fruit," "They are both necessary for life"), 1 point to a shared specific attribute or shared function (e.g., "They both have wheels," "They are both used to travel"), and no points to irrelevant responses or the failure to respond at all. It can be seen readily that these scores (with the exception of the reversal of groups 1 and 3) generally parallel the proportions reported in column 1 and are more amenable to statistical analysis. Consequently, these WAIS scores were subjected to simple analysis of variance with the result that there was an overall reliable groups effect, $F(5,66) = 3.13$, $p < .05$. Contrasts seen using the Newman-Keuls procedure with $\alpha = .05$ indicated that the only reliable differences that existed were between the secondary school students and the more highly educated Mestizo adults on the one hand, and the third-grade sample. All other comparisons were nonreliable.

Turning our attention to those columns of table 9 representing subjects' justifications of their correct choices, we see the same kind of marked differences between groups we observed in the previous study: semantic justifications increase both as a function of age and education.

If we consider the first three populations (i.e., third and sixth graders plus Mayan noneducated adults) as a set (roughly 40% correct), we see that the younger students differ from the Mayan adults most strikingly in their use of attribute rather than functional justifications. In the last three populations (the two bilingual Mestizo adult populations plus the secondary students; roughly 50% correct), the education difference apparently shows up both as an increase in semantic justifications and a decrease in functional justifications.

CONCLUSIONS

These results confirm the influence of both age and education of verbal classification seen in the preceding experiment. They also confirm our conclusion that the estimated magnitude of the effect depends upon details of procedure and the aspect of the data we choose to look at.

From the experiments described thus far it is clear that all populations studied *can* make verbal classifications on the basis of shared semantic features. But as the task takes on extra properties (such as the need to ignore functional relations, or insistence that the semantic classification be based on a very general category), semantic attributes may not be used, especially by the less-educated subjects; instead, attributes of the objects, or their functions, come to guide subjects' responses. If correct responding can be achieved using a functional classification scheme, educational differences are reduced vis-à-vis the effect of age; even secondary school students will use functional explanations in cases where either a functional or semantic criterion produces a correct choice.

III. STUDIES OF MEMORY AND THE LEARNING OF CLASSIFICATIONS

All of the studies reported thus far have focused on the ways in which subjects classify or reclassify objects physically arrayed before them or presented a very few at a time. All population groups produced taxonomic groupings in the face of one or more of the demands represented by our various tasks: populations differed primarily in the range of tasks which they responded to in terms of latent taxonomic categories.

In the present section we will report the results of several studies in which classifying is not obviously demanded of the subject; rather, each study represents a developmentally sensitive cognitive task in which developmental differences are often ascribed to the more effective or pervasive use of category information by older subjects. Each of these studies has the additional characteristic that the information upon which the subject must act is presented in such a way that he must rely heavily on his memory in its various aspects in order to respond effectively. Differences in memory demands as well as the nature of the categorization scheme will be the foci of the experimental manipulations.

FREE ASSOCIATION

Originally thought of as a tool for understanding the laws of thought, and later as a probe of the unconscious, free associates have been used in the past 20 years as a technique for understanding the development of the lexicon.

There are two basic phenomena in research on free associates that have attracted the attention of developmental psychologists. When familiar words are presented to American college students ("Tell me the first word that comes to mind when I say 'house' "), the response is likely to be of the same grammatical class (in this case, a noun) and the same semantic class (home, cottage, building) (see Entwistle 1966; Moran 1966). The change from responding with a word which enters into a sentence with the stimulus word (grumpy-daddy) to responding with a word from the same grammatical form class (grumpy-awful) has been termed the "syntagmatic-paradigmatic shift." Both the tendency to give responses in the same grammatical form and the tendency to give semantically related responses increase with age, tested mental ability, and socioeconomic status (see Entwistle 1966; Rosenzweig 1964). In our own work in Liberia (see Cole et al. 1971; Sharp & Cole 1972), we found that both these response tendencies were associated with higher levels of education; at the lower secondary school level (grades 7–9), educated Kpelle children gave more responses matching the stimulus word both grammatically and semantically.

Of special interest in the present context is work on the development of associative responding reported by Entwistle (1966) and Rosenzweig (1964) in which *both* educational and cultural factors, in addition to increasing age per se, were implicated as important contributors to the "syntagmatic-paradigmatic" shift. Rosenzweig, comparing French workers and university students, found that the former gave less-paradigmatic responses, a result that could be attributed to either educational factors, cultural factors, or both. Similarly, Entwistle found that children from the Amish region of Eastern Pennsylvania, whose primarily agricultural culture prides itself on its separation from, and lack of dependence upon, American mainstream culture and its trappings, shifted from syntagmatic to paradigmatic responding later than children from Baltimore. In explaining these results, Entwistle attributed the Amish children's deficit to the lack of opportunity that the child raised in Amish culture had to interact with mid-twentieth-century American culture. Taking the argument one step further, she noted differences between her results and responses to the original Kent-Rosanoff (1910) norms for words that were common to both studies, again suggesting that cultural factors are an important determinant of changes in verbal associations.

It is not clear to us *why* education should affect associative responding nor the nature of the cultural factors that may be at work. One line of interpretation would be that increased educational experience increases paradigmatic responding indirectly via increases in vocabulary knowledge. The importance of vocabulary knowledge was paramount to Luria (1930/1974), who carried out one of the first comparative studies of word associates. He hypothesized that educational experience and adult-child interaction increase the child's actual vocabulary, enriching and strengthening the basic semantic relations among the words in his repertoire. Such an interpretation finds support in the work of Stolz and Tiffany (1972) who found that even college students will give "child-like" associates to words that are familiar but infrequently encountered. Hall (1972), who found that education increases paradigmatic responding, believed that emphasis on conceptual word usage in the lower grades is responsible for the effect. Whether the change represents differences in the development of important language skills or a set to respond in a paradigmatic fashion is not clear.

We conducted several free-association studies with populations varying in terms of age, education, and language. In addition to these population variations, we introduced variations in our instructions and stimulus words to help evaluate the notion that populations differ in their implicit understanding of, or reaction to, instructions.

Responses to Nouns, Verbs, and Adjectives with Age and Education Confounded

Our initial study was a simple replication of free-association procedures used previously in both the United States and Africa to determine the basic pattern of results for schoolchildren.

Subjects.—Ten subjects from each of the following populations were selected: (1) 6–8-year-old students (grade = 1, age = 7.3), (2) 9- and 10-year-old students (grade = 5, age = 9.6), (3) 11–14-year-old students (grade = 6, age = 11.9), (4) 14–17-year-old students (grade = 8.2, age = 15.9). The elementary school students were seen at school one at a time in a separate room provided by the teaching staff. The secondary students, on the other hand, were seen after school in their own homes or the home of one of our assistants.

Stimuli.—The stimuli were 18 common words from the three grammatical form classes: nouns (knife, plate, pot, bread, tortilla, and trousers, stimuli that should be familiar from previous studies), verbs presented in the infinitive form (to eat, to play, to buy, to carry, to run, and to climb), and adjectives (large, red, pretty, good, expensive, and tall).

Procedures.—All subjects were given the following instructions in Spanish:

> I have here a list of words that I will read to you one at a time. Each time that I read you a word, I want you to tell me five words that you might think of. For example, if I hear the word "house," I think of building, post office, store, shoe shop, tortilla shop, as well as "house." Or if I hear the word "boy," I think of girl, woman, man, and so on. Do you understand what I want?

Subjects generally expressed little difficulty understanding what was required of them. A different fixed order of presentation was used for each subject such that a verb followed a noun and was in turn followed by an adjective in the order noun-verb-adjective throughout the list. Responses to each stimulus word were written down in the order given. (It should be noted that, although five response words were required of each subject for each stimulus word, some subjects, particularly in the younger age groups, failed to come up with the required number and were allowed, at the discretion of the experimenter, to stop if they showed undue hesitation or complained that they could think of no more.)

Results.—Responses were scored to determine the proportion of the same form class (paradigmatic responses) for each of the three form classes for each subject. The resulting proportions were transformed using the arcsine transformation recommended by Winer (1962) prior to a groups × form class analysis of variance with repeated measures on form class.

The results of this experiment are presented graphically in figure 4. The analysis of variance confirmed the trends that are displayed in figure 4 in that both the main effects and their interaction were highly reliable—F(form class)(2,72) = 147.73, $p < .01$; F(groups)(3,72) = 7.15, $p < .01$; F(groups × form class)(6,72) = 2.81, $.01 < p < .05$). Individual comparisons run using the Newman-Keuls procedure with $\alpha = .01$ indicated that the only reliable groups difference was between the secondary school students and the other three groups taken as a unit. The differences between the three form classes were all highly reliable with the proportion of paradigmatic responses to nouns being reliably higher than that to adjectives, which in turn were reliably higher than the proportion to verbs, a result that parallels findings reported from both the United States and Western Europe.

The significant interaction of groups × form class, as figure 4 makes clear, demonstrates that the difference between groups is smallest when nouns are used as stimuli. This result has been obtained repeatedly, both in research in the United States and abroad (see Hall 1972), but its origin remains obscure. One important implication of the result, however, is to emphasize that the failure of the younger (less-educated) groups to respond

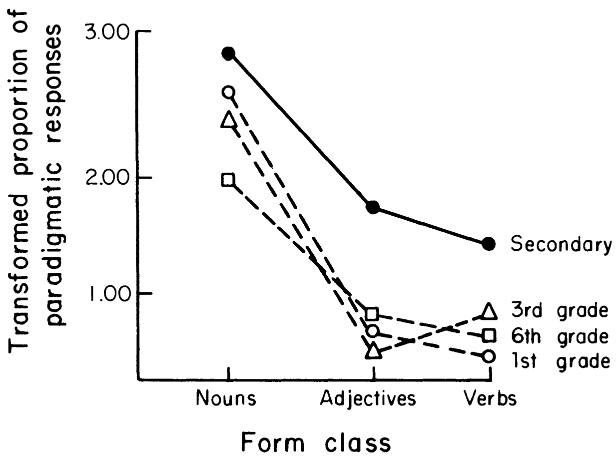

FIGURE 4.—Paradigmatic responding as a function of form class for students

with words of the same form class is not a general characteristic of their total vocabulary but related to the form class of the word presented (and, as we shall see, the instructional set induced by the experimenter).

Age versus Education

We repeated the previous study, dropping two of the younger groups and adding in their place two adult groups. Members of the two adult groups were nonliterate but represented the two major subcultural groups of rural Yucatan, Maya and Mestizo.

Subjects.—The subjects, all from Ticul, 12 to a group, were drawn from the following four populations: (1) 9–10-year-old students (grade = 3, age = 9.4); (2) secondary school students (grade = 7.7, age = 15.8); (3) 20–40-year-old nonliterate, bilingual Mestizo adults (grade = 0, age = 29.2); (4) 20–40-year-old monolingual, nonliterate Maya adults (grade = 0, age = 31.6).

Two of these groups overlap the groups employed in study 1, the third graders and the secondary school students, a fact that allows us to make comparisons between the two experiments. In addition, the presence of the two nonliterate adult populations allows us to check for age effects independent of educational experience and for possible language or cultural effects in the absence of education. The two nonliterate adult groups are of comparable age and educational experience, but Mestizos as a group are generally considered to be more acculturated with reference to mainstream Mexican culture than the strictly Mayan population.

Stimuli and procedures.—The stimuli and procedures were the same as employed in the prior study with the exception that the monolingual Mayan adults were instructed in, and asked to respond in, Maya, while the bilingual, Mestizo adults were instructed in, and responded in, the language with which they felt most comfortable.

Results.—To begin with, we ran t tests comparing the groups that overlapped between studies 1 and 2; they were nonsignificant, $t(22) \leq 1.00$, $p > .10$ in each case, indicating comparability between the two studies. The results of this study are presented in figure 5.

The major results are apparent from the figure. All groups are more likely to give paradigmatic responses to nouns than verbs or adjectives, $F(2,88) = 68.09$, $p < .01$. Similarly, the main effect for groups was also highly reliable, $F(3,44) = 13.98$, $p < .01$, as was the groups $\times$ form class interaction, $F(6,88) = 7.53$, $p < .01$. Individual comparisons again were run for this groups comparison using the Newman-Keuls procedure with $\alpha = .01$, with the result that the secondary students differed reliably from the two adult groups who were statistically equivalent and in turn differed reliably from the third-grade sample. Since the two adult groups have experienced no education and are older than the secondary school students,

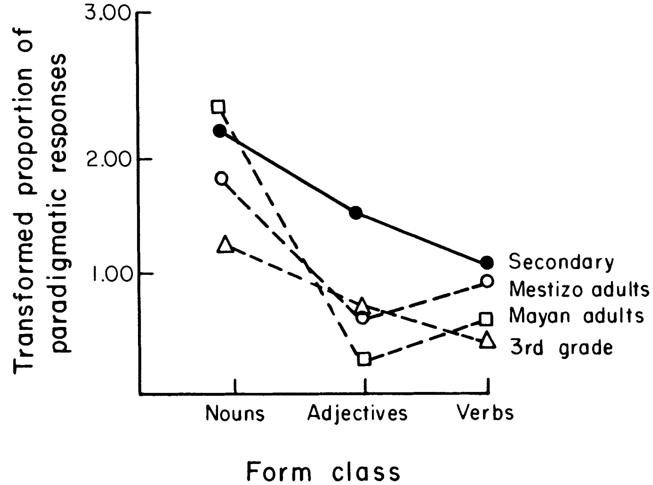

FIGURE 5.—Paradigmatic responding as a function of form class for different population groups varying in age and education.

it would appear clear that both increases in age and education make independent contributions to the incidence of paradigmatic responding.

Form Class and Semantic Class

At the beginning of this section on free association we emphasized that two kinds of classification were involved—according to grammatical form class and semantic class.

Studies 1 and 2 established the existence of independent age and educational differences in the tendency to respond paradigmatically.

In this study, we will use the strong tendency of all groups to give noun responses to noun stimuli in order to investigate the second kind of classification that is discernible in free-associate responding—responding with semantically related words.

Subjects.—Twenty subjects were chosen from each of the following five population groups: (1) 8–10-year-old children from Mayan households (grade = 2.0, age = 9.7), (2) 8–10-year-old children from Mestizo households (grade = 2.4, age = 8.5), (3) Maya adults (grade = 1.2, age = 23.6), (4) Mestizo adults (grade = 3.3, age = 20.8), (5) secondary school students (grade = 8.2, age = 16.9).

Stimuli.—Twelve common nouns from three of the semantic classes used in our picture classification study were used: clothing (dress, sandals, shirt, trousers), food (onion, tortilla, coconut, bread), animals (bull, horse, duck, dog).

Procedure.—Although our major concern in this study was to obtain data on population differences in the use of semantic relations between stimulus and response words, an additional issue caused us to vary instructions within each population group. In particular, we wanted to see if we could influence the form class of the response by varying the nature of the examples we gave as part of our instructions.

Half (10) of the subjects in each group were given instructions identical to those in studies 1 and 2. For the remaining subjects the examples were varied, so that (e.g.) in response to the hypothetical stimulus word "building," subjects were given hypothetical responses, as "red," "to live," and "store," instead of nouns from the same semantic class ("store," "house," "hotel"). In all other respects, the procedures were identical to each other and to the procedures used in studies 1 and 2.

We suspected on the basis of pilot work that varying the form class of the example responses would decrease paradigmatic responses to nouns. We were particularly interested in finding out if educated subjects were more sensitive to the instructional variation than their less-educated counterparts.

Results.—All groups were influenced by the instructions: nouns predominated when noun examples were given, while adjectives predominated when a mixture of all three grammatical forms were given. Moreover, the change between instructional conditions was greatest for the secondary school students. These results demonstrate that not only fixed aspects of subjects' vocabularies but also the example responses given as part of the instructions influence subjects' choice of responses. Since we did not run a full factorial experiment varying instructional set and word class, these results should be taken as no more than a demonstration of potential interactions that need further, systematic research.

The main point of this study was to determine the extent to which responses to noun stimuli would be from the same semantic as well as grammatical form class.

Table 10 contains three measures of the cohesion and semantic organization of subjects' responses to noun stimuli.

The top row in table 10 ("Group communality") reflects the extent to which different subjects respond with a small set of words, regardless of

TABLE 10

EVIDENCE OF SEMANTICALLY ORGANIZED ASSOCIATIONS TO NOUNS

	Mayan[a]	Mayan Adult	Mestizo[a]	Mestizo Adult	Secondary School
Group communality[b]	.27	.23	.29	.29	.32
Semantic responses	8.00	9.70	6.40	9.80	35.40
Superordinate responses	1.20	2.40	1.50	2.50	10.0

[a] 8–10 years old.

[b] Communality is taken as the ratio of the four most frequently given responses to the total number of responses to each stimulus word. The data are averaged across all 12 stimulus words.

semantic class. Analysis of variance run on the data yielded an overall $F(4,55) = 5.99$, $p < .01$; this result and appropriate contrasts using the Neuman-Keuls procedure support the conclusion that the secondary school students show greater communality than the remaining groups, while the Mayan adults show the least. The differences, however, were not substantial. This pattern of results is consistent with increasing grade as a causal factor in the structuring of associative responding because the Mayan adults are least educated of all the groups, while the secondary school students are the most educated. The secondary school students also give vastly more semantically related words, $F(4,55) = 68.14$, $p < .01$, than the remaining groups (where individual group contrasts indicated that the remaining groups did not differ from each other). Finally, in 10 out of a possible 12 cases, the secondary school students also gave a superordinate term as one of the semantically related terms. Here again, we see an apparent effect of grade but only when high levels are contrasted with 1–3 years.

Conclusion.—The semantic association results of study 3 stand in marked contrast to the form class data from the previous two studies. Whereas there is clear evidence for the expression of *grammatical* form class responding to nouns in our younger and less-educated samples, the nouns chosen are not generally of the same semantic class as the stimulus word. Only in the case of secondary school students do we see both a generalized tendency to respond paradigmatically *and* a very consistent choice of semantically based selection within the domain of nouns.

We can view the subject's task in these studies as one of selecting "appropriate" responses from long-term memory store to match with a single stimulus item. Education, age (or some experiential concomitants of age), and the particulars of the task differently affect both subjects' understanding of what is appropriate and what to do given this understanding.

We are fully aware of the many unresolved issues raised by these studies. It would be useful, for example, to explore more deeply the effects of instructional variations as they interact with the stimulus set. However, we did not pursue these problems because of limitations on time, combined with our conviction that the major results are best interpreted in the framework of related studies to which we now turn. An interpretive summary may be found in Chapter VI.

FREE RECALL

The "remembering" aspect of the free-association task consists almost entirely of retrieving lexical items from long-term memory; there are no correct or incorrect responses in any general sense.

Our previous work in Liberia was consistent with the current free-association results in finding that semantically controlled associations to nouns apparently became dominant only when subjects had experienced some level of secondary school education.

We also found that secondary school was the level at which semantic organization appeared in a task that more directly required memorizing—free recall.

In current theoretical terms (Murdock 1974), the free-recall task contains a mixture of short-term and long-term memory requirements. Subjects are required to remember a set of items (often words read aloud) that is too long to retain in short-term memory. In the form of this task to be used here, the items are representative of a small set of semantic classes. The items and their category structure are typically well known to the subject as part of his general language knowledge (although he may not recognize or use the category relations in the list when recalling it).

The present free-recall study was conducted with adults of varying ages and degrees of educational experience. It was preceded by a pilot study with schoolchildren in grades 2–9 who varied in age from 7 to 17 years. The pilot work had shown, consistent with our previous experience, that there were small developmental differences between grades 2 and 5, with a significant increase in both recall and semantic clustering appearing for secondary school students.

In light of the free-association data regarding semantically controlled responding and these pilot free-recall results, we expected a significant effect of education on free-recall performance, with little or no effect of age across the age span from 18 to 40 years.

Subjects.—Subjects, roughly two-thirds of whom were male, were drawn 12 to a group from the following five subject populations: (1) nonliterate Mayan adults (age = 58.0), (2) nonliterate Mayan adults (age = 27.7), (3) bilingual Mestizo adults with 1–3 years of formal education (grade = 2.1, age = 28.7), (4) bilingual Mestizo adults with 4–6 years of formal education (grade = 4.9, age = 28.1), (5) educated adults: ex-students and students with 7 or more years of formal education (grade = 9.9, age = 21.1).

This design allowed for both age and education comparisons within an adult sample. For example, the two educated Mestizo groups are roughly the same age but are significantly older and significantly less educated than the educated adults. Similarly, the two Mayan groups allow for age comparisons in a noneducated, adult population.

Stimulus materials and procedure.—The items to be remembered were the 20 common nouns used in the picture classification study: animals (bull, dog, duck, horse, and turkey), utensils (pot, plate, bottle, cup, knife), food (tortilla, coconut, corn, onion, bread), and clothing (shirt, huipil, pants, sandals, and hat). Subjects were told that they were going to be told the names of several things which they should try to remember so that they could name them all. The list of 20 words was read at a rate of about one item every 2 seconds in the language in which the subject felt most comfortable. After presentation of the list, the subject was asked to recall what he had heard in any order he wished. The list was presented five times, each in a different random order, with the restriction that no two items from the same category occur adjacently.

Results.—The data were analyzed in three ways. First we will consider simply the number correctly recalled. A repeated-trials analysis of variance revealed that the overall difference between groups was significant, $F(4,55) = 4.91$, $p < .05$. Those with secondary school experience recalled an average of 12.6 items, the fourth- to sixth-grade subjects recalled an average of 12.4, and the remaining groups averaged about 10.8 items per trial. All groups improved from trials 1 to 5, $F(4,220) = 80.71$, $p < .01$. Contrasts indicated the two more highly educated groups were statistically equivalent and superior to the three remaining groups.

The results in terms of semantic category clustering were more dramatic and are most readily interpreted by reference to figure 6.

Our measure of semantic category clustering is the Z score first suggested by Hudson and Dunn (1969) and used extensively in our earlier work.[5]

From the figure it is clear that extensive category clustering is to be found only among subjects who had experienced a secondary school education and that even for them clustering is present only on later trials. The remaining groups perform at approximately a chance level of organization.

This free-recall study is one of the few experiments in which we also obtained data for each subject concerning demographic variables such as occupation, the amount of schooling available in his hometown, family background, etc.

The major conclusion from multiple-regression analyses run on the recall and clustering data is that grade is the key variable associated with improved performance. The only variable which would enter significantly into an equation with grade was speaking Spanish as a first language, which exerted a small negative effect. Other variables such as parental occupation and education, age, sex, and language either have no effect at all (as in the case of sex, e.g.), or work indirectly through their effect on grade.

An additional aspect of the results which bears noting in light of results to be reported shortly is that *all* the groups in this study manifested a significant serial position effect; the first two items and last three items presented on each trial were recalled better than items presented in the middle of the list.

This observation is relevant to our comment at the outset of this section

[5] The Z score is based on the distribution of run lengths, providing a test of the difference between observed and expected average run lengths of items within a specific semantic category.

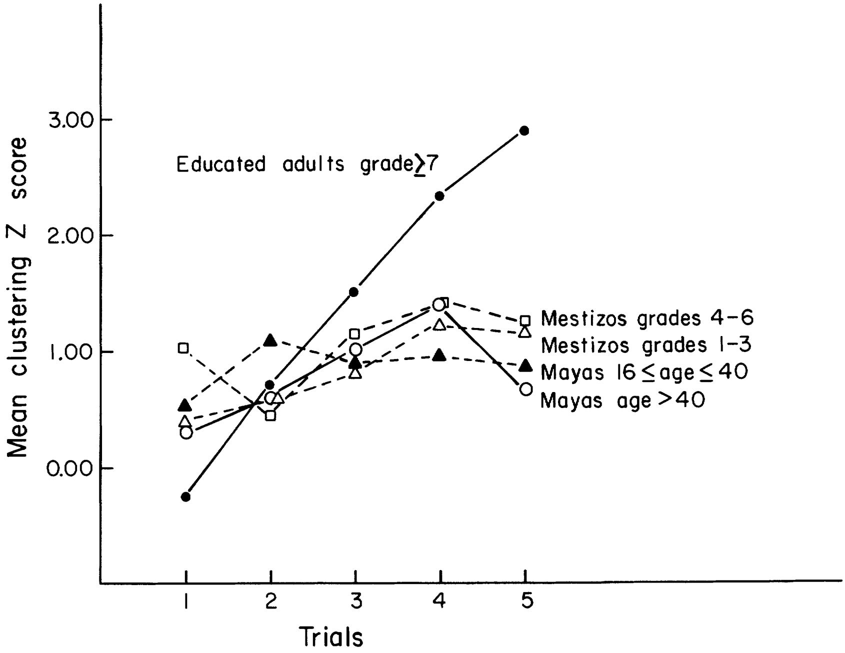

FIGURE 6.—Organization of free recall for different population groups varying in age and education.

that free recall is generally viewed as a mixture of short-term and long-term memory processes. Whatever that mix, it appears constant across groups for this task; the serial position effect was of roughly the same magnitude for all groups (the groups × serial position interaction was not significant; $F = 1.1$).

PAIRED-ASSOCIATE LEARNING

On the surface, at least, the results of the classification studies and the results of the studies where classification has a role to play in memory tasks present us with a contradiction. In our initial object-sorting study and in parts of other classification tasks, nonliterate subjects, both children and adults, manifested knowledge of semantic categories which they used in responding to our instructions. In the studies where remembering was a central component, however, only more highly educated subjects used category information in dealing with the task.

The present study pursues the hypotheses that the less educated *can* use semantic information to guide their responding in a memory task but will do so only when the utility of doing so is "built in" to the task.

Our vehicle for testing this notion is a paired-associate experiment in which the items to be associated are selected according to two different principles. In one case (within-category), the items to be associated are selected from the same semantic class (bull-sheep); in the second case (between-category), items are from different semantic classes (bull-root).

From our sorting studies and well-documented effects of interitem associations on paired-associate learning (see Murdock 1974), we expected the within-category condition to be easier than the between-category condition. However, the free-associate and free-recall data make this conclusion problematic. We particularly wanted to see if the presumed facilitative effects of the within-category condition were equivalent for different age-grade groups.

Subjects.—Twenty subjects were selected from each of the following subject populations; roughly half of each group (53 out of the total 100 subjects) were male: (1) monolingual Mayan adults (age = 39.4), (2) bilingual Mestizo adults (grade = 4.8, age = 27.7), (3) third-grade students (age = 9.1), (4) sixth-grade students (age = 14.0), (5) secondary school students (grade = 8.3, age = 15.0).

Stimuli and procedures.—Each subject was presented one of the following two 14-item paired-associate lists:

WITHIN-CATEGORY PAIRS	BETWEEN-CATEGORY PAIRS
root-branch	root-sheep
plate-bowl	dress-hill
chicken coop–market	plate-crocodile
dress-shawl	breath-bowl
calabash-comal	calabash-town(people)
breath-thought	island-branch
coconut-papaya	well-machete
strength-life	strength-market
hoe-machete	brush(woods, jungle)-comal
island-town	hoe-cave
frog-crocodile	coconut-shawl
brush-hill	frog-thought
bull-sheep	chicken coop–life
well-cave	bull-papaya

The following instructions were given in Spanish (for the school-attending groups and the Mestizo adults) or in Maya (for the Mayan adults):

> I have here a list of words that I will read to you in pairs. The first time that I read the list, I want you to repeat each pair. Then I will read the list again mentioning only the first word of each pair, and I want you to tell me the second word of each pair. I am going to give you 5 seconds for you to try and remember the second word of each pair. Do you understand what I want you to do?

Subjects generally had no trouble understanding what was required. Students were seen in a separate room provided by the administration of their respective schools, while adults were generally visited in their own homes during the evening. Subjects were presented the list 10 times, or until they correctly anticipated all 14 response terms correctly.

Results.—The major results of this study are contained in figure 7. Performance was scored in terms of the total number of errors committed up to one completely correct recall of the list or in 10 trials (subjects who recalled all items on a trial perfectly were assumed to have no further errors). An inspection of the graph represented in figure 7 makes it clear that both main effects and their interaction were highly reliable. The within-category list was substantially easier than the between-category list, $F(1,90) = 74.7$, $p < .01$; the groups order themselves almost perfectly in terms of the number of grades they had completed, $F(4,90) = 16.1$, $p < .01$; and, finally, this grade effect appears most clearly marked for the between-category condition, the F ratio for the groups $\times$ conditions interaction $(4,90) = 3.8$, $p < .01$.

Conclusion.—This study demonstrates that, if the task is appropriately structured, subject populations with low levels of education can benefit from knowledge of the category structure inherent in the to-be-remembered materials. A major difference between the paired-associate experiment and the previous studies was that in the present case the relationships were built into the task (in the case of the within-category condition), whereas in the

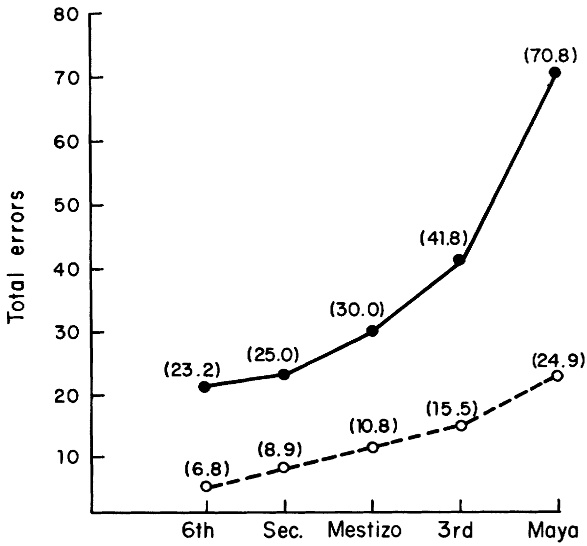

FIGURE 7.—Total errors to criterion in paired-associate learning of population groups varying in age and education as a function of item-category membership relations. The solid line indicates between-category pairs. The dotted line indicates within-category pairs.

free-recall and free-association studies the semantic information, while available, need not have been used. It may be significant also that this task requires that the subjects use relations between pairs of objects, while clustering in free recall requires coordination of five items in four categories.

Granting ambiguities arising from differences in the stimuli used in these different memory tasks, the results suggest to us that *educational effects will appear most strongly in those cases where the task allows, but does not require, the subject to apply his semantic knowledge.*

On the basis of similar results from our research in Liberia and in the United States (see Brown 1975b), we can speculate that the use of semantic knowledge in the free-recall and free-association studies is a special case of the more general phenomenon that sophisticated subjects actively restructure the task as presented in the interests of enhanced performance.

SHORT-TERM RECALL OF LOCATION

If our speculations based upon the results of the studies reported thus far are correct, we ought to observe particularly strong grade effects in those tasks (or parts of tasks) which permit, but do not require, the subject to supplement the raw information presented by engaging in some cognitive activity or activities (such as the use of category knowledge in free recall) which produces enhanced performance.

Exactly such an effect was obtained in a study by Wagner (1974), who was working in Yucatan at the time our studies were being conducted. Since Wagner's study is published, we will present only those highlights important to our discussion.

Wagner conducted a simple and easy-to-administer short-term memory task, first introduced by Maccoby and Hagen (1965). Using pictures of animals and common domestic objects taken from the popular Mexican game of lotteria, similar in nature to American bingo, Wagner constructed sets of stimuli on 1×3-inch to $1\frac{1}{2} \times 3$-inch cards where each card contained a picture of an animal and a picture of a domestic object. A given set contained seven animal-object pairs and a "probe" card which contained a picture of an animal or an object. The basic task was for the subject to remember the position of an animal (for half the subjects in each group) or an object (for the remaining half) after the cards had been laid out in a linear arrangement with 2 seconds for observation of the position of each card (see Maccoby & Hagen [1965] for the prototype of this procedure). Probe trials were used to test recall of items in a random order in each of the seven possible positions. Since the cards were always laid out from the subject's left to his right, the position of the probe also determined the delay between presentation of a card and the text for recall of the position of the card.

The subjects in Wagner's study were selected from five different age ranges: 7–9 years, 10–12 years, 13–16 years, 20–21 years, and 22–35 years. Education varied from 0 to 12 years of schooling. Unlike the other studies thus far reported, Wagner attempted a systematic investigation of the influence of urbanization on the development of cognitive performance by selecting roughly half of the subjects from the small, relatively traditional Mayan town of Mayapan and the other half from the capital city of Merida. The Meridians, except for the two youngest age levels, had experienced considerably more years of schooling than the rural subjects.

The major result is shown in figure 8, which plots the proportion of correct recall responses as a function of age and urbanization. Statistical analysis revealed that only after educational levels diverged (age 13–16 years) did the performance between urban and rural groups differ, although the difference that did exist was in the predicted direction. Thus, it appears that what Wagner observed was basically a difference in memory performance owing to the increased efficiency of the more-educated subjects and the lack of improvement on the part of older subjects who had not been to school.

Critical to our current argument were the results of Wagner's analysis

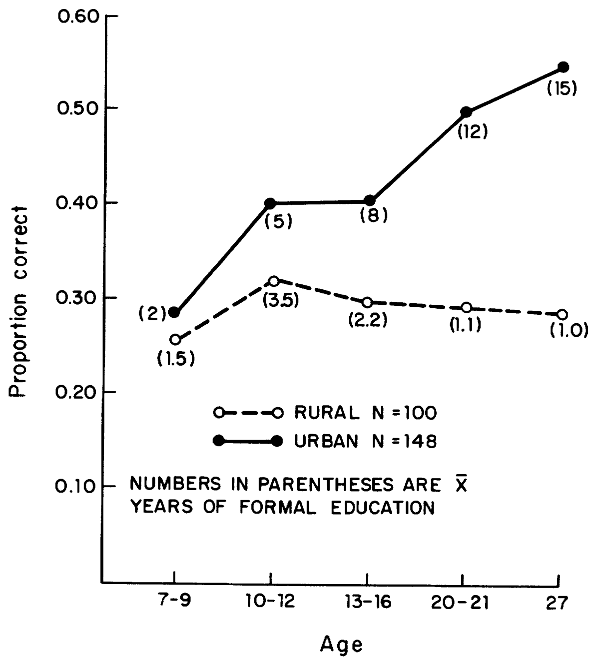

FIGURE 8.—Short-term recall of location as a function of age, education, and urbanization (after Wagner 1974).

of response accuracy as a function of the time delay between stimulus presentation and recall, for example, the short-term and (relatively) long-term portions of recall. He found for the item presented just prior to recall ("short-term") there were no differences in performance between groups. The difference plotted in figure 8 arose entirely from the enhanced performance of the more-educated subjects on the early and middle ("long-term") items of the stimulus set.

On the basis of extensive research using this probe recall paradigm with children in the United States, Wagner concludes that the performance differences he found must be attributed to spontaneous application of remembering strategies (especially rehearsal) by the more-educated subjects. The lack of difference between groups on the item presented just prior to recall fits this interpretation neatly because the last item is assumed to be more or less automatically retrievable from "echoic" memory (see Hagen & Hale 1973).

Wagner's conclusion also dovetails with our previous findings and generalizes them in a very useful way. Not just the importation of semantic information (our point of departure in the previous studies) but the active application of more general procedures or strategies for bringing order to an experimental problem seems to be a general factor underlying improvements in performance as a function of education.

IV. STUDIES OF PROBLEM SOLVING

In an important sense, all of the tasks presented in previous sections require the subject to transform information presented to them in order to comply with instructions; in that limited sense, all are problem-solving tasks.

The two studies presented in this section are organized under the heading of problem solving partly because they include tasks which are typically labeled in this way in the psychological literature. A more compelling reason is that most of the tasks presented here are intended to focus on the way information is (or is not) transformed to produce an answer to an explicitly posed problem. Memory and classification are understood to be important components of this process, and memory tasks are included along with problem-solving ones in the second study.

VERBAL LOGICAL PROBLEMS

Research in quite different parts of the world (Cole et al. 1971; Luria 1976) indicates that relatively low levels of education influence subjects' responses to verbal logical problems such as syllogisms. While the basis for this change is by no means clear (see Scribner 1976b), it appears that the educated subjects more readily accept the problem as self-contained information to be evaluated in its own right, while less-educated subjects are likely to "assimilate" the problem to their own past experience. The present study follows the basic design of the previous studies in this report in an attempt to determine the influence that educational or other experiences have on responses to verbal logical problems. It is the only experimental study (with the exception of Wagner's short-term memory task) in which groups from two towns varying in sophistication or degree of acculturation are included as a part of the design.

Subjects.—Subjects were drawn from the following subject populations (number of subjects per group indicated separately, as is the town where they are living; 57% of the subjects in this study were male): (1) first- and second-grade students from Ticul (grade = 1.8, age = 8.8, N = 33), (2) nonliterate or near nonliterate Mayan adults from Ramonal (grade = 0.7, age = 32.1, N = 32), (3) nonliterate Mayan adults from Ticul (grade = 0, age = 48.0, N = 41), (4) Mestizo adults from Ticul (grade = 3.5, age = 34.1, N = 36), (5) Mestizo adults from Ticul (grade = 4.2, age = 36.5, N = 30), (6) fourth- to sixth-grade students from Ticul (grade = 4.8, age = 12.4, N = 30), (7) secondary school students from Ticul (grade = 8.7, age = 15.9, N = 30).

Ticul is one of the largest and most sophisticated towns in our study, while Ramonal is a typical small, relatively traditional, Mayan village (see fig. 1 for a partial comparison of their characteristics).

Stimulus materials and procedures.—The verbal logical problems we used were adapted from previous research in Liberia (see Cole et al. 1971, chap. 6). The problems are listed in table 11.

The subject was read the problems one at a time in either Spanish or Maya (whichever language was more convenient for the subject). He was required first to answer the question and then to explain the reason for his answer.

Results.—The results summarized in table 12 are evaluated both in terms of the proportion of correct responses given by each population group and in terms of the reasons given for the choice of response.

TABLE 11

Problems for Logical Problems Study

1. A dog and a horse are always together; the horse is here now; where do you think the dog might be now?
2. Where the earth is good crops are good; the earth in Kilometro 71 is good; do you think the crops are good?
3. Women from Guadalajara are pretty; my (female) friend is from Guadalajara; do you believe my friend is pretty?
4. The houses in Mexico City are large; my friend has his house in Mexico City; how do you think that his house might be?
5. If Juan and José drink a lot of beer, the mayor of the town is angry; Juan and José are drinking a lot of beer now; do you think the mayor is angry with them?
6. So that José might be able to carry corn from his farm to the town center he needs a bag and a horse; he has the horse, but doesn't have the cart; can José carry the corn from his farm?
7. In a chicken coop if a hen is eating or a rooster is eating, the turkey wants to eat; the hen is eating, but the rooster is not eating; do you think the turkey wants to eat?
8. If in a chicken coop a hen and a rooster are eating together, the turkey wants to eat; the hen is eating now, but the rooster is not eating; do you think that the turkey wants to eat?
9. If Juan and José don't drink a lot of beer the mayor is angry with them. Juan is drinking a lot of beer now, but José is not drinking; do you think the mayor is angry?
10. In order to work well one has to give a horse enough to eat; José gives his horse enough to eat; do you think José's horse works well?

TABLE 12

RESPONSES TO VERBAL LOGICAL PROBLEMS

	Ramonal Mayan Adults	Students, Grades 1 and 2	Ticul Mayan Adults	Mestizo Adults (Grade 3.5)	Mestizo Adults (Grade 4.7)	Students, Grades 4–6	Students, Secondary
1. % correct responses	45	53	63	73	76	78	97
2. % theoretic justification	43	48	59	73	79	82	99
3. % correct theoretic	98	99	99	100	100	98	100
4. % correct empiric	64	65	63	65	67	50	[a]

[a] No data.

Taking problem 3 as an example, we might imagine the following interchange:

EXPERIMENTER: Women from Guadalajara are pretty. My friend [female] is from Guadalajara. Is my friend pretty?

SUBJECT: Yes, she is pretty. You like to be with pretty women, so you will choose pretty women no matter where they are from.

The answer is correct and is so scored. But the justification is in terms of personal or normative knowledge, not a response to the logical relations given in the problem. Following Scribner (1976b), we shall term these "empiric" responses. If the subject had answered to the effect that "yes, you said all women from Guadalajara are pretty," we accepted the form of the reason as evidence that the subject was responding to information contained within the problem. We shall term these "theoretic" responses.

A simple groups × treatments analysis of variance on number of problems correct yielded a highly significant overall $F(6,225)$ for groups of 25.37, $p < .01$. Individual contrasts employing the Newman-Keuls procedure with $\alpha = .01$ supported the conclusion that the secondary students performed better than any of the groups that had received at least a third-grade education but less than sixth grade on the average (i.e., the two Mestizo groups and the fourth to sixth graders from Ticul); these groups, in turn, perform better than the first- and second-grade children or the Mayan adults. Some experience other than education, however, must influence responses in this kind of situation because the Mayan adults from Ticul respond significantly better than a comparable population from the smaller, more traditional town of Ramonal. It is difficult to say what possible experiential variable is most likely to account for this result, but the fact that most Mayans in Ramonal are campesinos engaged in agricultural tasks while many of the Mayans from Ticul may be engaged at least part-time in some sort of trade may well be important, not to mention the fact that the towns differed markedly at the time of our studies in their contacts with the outside world and the services that they provided.

Several further analyses were done on these data to assess the importance of education within the adult populations which had experienced at least some education. Correlations and partial correlations were run between age, number of correct choices, and grade. In each case, there was a statistically significant positive correlation between grade and number correct. The correlations and partial correlations between age and number correct were never significant in the positive direction, although in one case (the Mayan adults from Ramonal) there was a significant *negative* correlation between age and number correct when grade was partialed out. This result almost certainly reflects the greater traditionalism of the older of the subjects in this group who are from a more traditional Mayan town.

We also conducted regression analyses in which the effects of sex, home language, and town modernity were evaluated. As suggested in table 12, town modernity positively affected performance. Not evident in table 12 is that females performed slightly less well than males, with grade and other variables held constant.

Earlier we pointed out that populations may differ because the less-educated subjects tend to assimilate the problem to their own past experience. Following procedures worked out by Scribner (1976a), we examined the justifications that subjects gave for their answers to the syllogisms. If subjects explicitly related conclusions to the premises of the problem, they were given credit for a *theoretic* response; if the subject referred to evidence of what he knew or believed to be true, the response was scored as *empiric*.

Inspection of row 2 in table 12 readily demonstrates that the groups differ markedly. The secondary students almost always respond theoretically, followed by the groups with 4 or more years of education. Grade, however, is not the only important factor. The Mayan subjects from Ticul respond with theoretic justifications more than the comparable group from Ramonal.

The critical role of theoretic responding is clear from row 3 of table 12. All groups are always correct if they respond in the theoretic mode; empirical reasonableness controls responding in the remaining cases. In this regard, it is instructive to examine the content of the syllogisms where the two groups of subjects do not differ markedly (problems 5, 6, and 10). These turn out upon inspection to be precisely the problems on which a correct answer is completely in tune with the experience of the subjects, that is, mayors get angry with drunks, horses need food to work, and some sort of container to carry corn from the farm. But the beauty of women from Guadalajara, houses in Mexico City, and the farmland at Kilometro-71 are other things altogether; here the subject is unlikely to have any direct experience upon which to base his response.

THURSTONE'S PRIMARY MENTAL ABILITIES: A WITHIN-SUBJECTS STUDY OF AGE AND EDUCATIONAL VARIATIONS IN COGNITIVE PERFORMANCE

During the period in which the previous studies were being conducted, two colleagues from Harvard University, Jeremy Anglin and Joy Skon, designed a study based upon an adaptation of Thurstone's Primary Mental Abilities (PMA) test (Thurstone 1938).[6] The Thurstone subtests were chosen because they sample widely from a number of areas of intellectual activity and in many instances overlap experimental studies which are

[6] We wish to thank Mr. Anglin and Ms. Skon for permission to describe their study here.

carried out with independent groups of subjects in all our other work. In addition, Anglin and Skon also took a somewhat different approach to subject selection. Working primarily in Ticul (although some of the subjects were selected from Bacalar, 250 kilometers to the southeast), they selected pairs of subjects from the same family—either siblings or cousins. These pairs were selected to provide a range of educational experiences from a few years of primary school to completion of "preparatoria" and a range of *differences* in educational experiences within families from no difference to several years' difference in educational experience.

Subjects.—Pairs of subjects from 17 families participated in this study. They ranged in age from 11 to 25 years and in educational experience from 2 to 12 years. Thirteen of the 34 subjects were female.

Materials.—Nine different subtests were adapted from Thurstone's (1938) monograph on primary mental abilities for use with Mestizo children and young adults. The test names and a brief description of each are included in the following:

1. Verbal analogies. Two words are connected by a certain relation. The first word of another pair is presented, and the subject's task is to choose one of five possible comparison words which completes a verbal analogy equivalent to the example pair.

2. Arithmetic. The subject is asked to add, subtract, and multiply using paper and pencil if he (she) wishes.

3. Completion of a number series. Three numbers are given in series, and the subject must complete the series in a manner that is consistent with the principle that generated the first three members.

4. Reasoning. The subject is given a complete syllogism and is asked to judge whether or not the conclusions follow from the premises.

5. Perceptual analogies. Each problem consists of eight figures. The first three are called A, B, and C, and the next are numbered 1–5. The subject's task is to pick, from the five comparison figures, the one to place with C that will create an analogy such that A:B as C:(1–5).

6. Arithmetic reasoning. Subjects are presented five verbal problems involving arithmetic calculations that must be done without paper and pencil.

7. Vocabulary. A target word was presented along with five comparison words. The subject's task was to select the comparison word closest in meaning.

8. Memory. There are two memory subtests. First, a set of 20 nonsense figures was presented for inspection and the subject had to identify the 20 from a set of 60 figures including the original 20 and 40 distractor items. The second memory task was, in effect, a single-trial paired-associate test with two-digit numbers as stimuli and responses. After studying the original stimulus-response pairs, the subject was presented the stimuli and asked to provide the responses.

9. Figure classification. The subject was shown pairs of panels each containing several figures. Some rule (as, e.g., all the figures in the first panel are oriented vertically, all those in the second panel horizontally) distinguished the panel. Using figures in a third panel which contained figures selected according to the principles used in the first two panels in a scrambled arrangement, the subject had to indicate his understanding of the rule which distinguished the panels.

10. Twelve items from the Raven matrices test.

Procedures.—In general, the procedures followed those recommended by Thurstone in his presentation of the tests and results (see Thurstone 1938, pp. 22–57). At the start of the testing session each subject was told about the general contents of the tests. The actual tests were administered by a trilingual Mestizo experimenter. Skon checked to insure the appropriate presentation of the materials, observed and recorded subjects' responses.

Following Thurstone, an example of the procedures and a correct answer (or appropriate performance) were given for each test. There were also extensive verbal explanations and provisions for asking questions. Each item was scored using procedures outlined by Thurstone. A composite score for each test as well as a composite score over tests (a total PMA score) were calculated for each of the subjects.

Results.—Owing to experimental error, data were not collected on one subtest of the PMA from six subjects and two subtests from two subjects. Only one subtest was missing data from as many as three subjects. To render summary data comparable, all subjects were assigned a total PMA score that represents their score as an average of all the subtests to which they contributed data.

The arrangement of testing conditions and the nature of the data rendered the data most easily analyzable in terms of correlational analysis rather than analysis of variance. Consequently, a stepwise multiple-regression analysis with age, sex, and grade as independent variables was run on the total PMA score and each subtest for which there were data from all subjects.

The results of this analysis were particularly clear-cut. For the overall PMA score the correlation between grade and performance was .75. Neither age nor sex correlated significantly with overall PMA performance, and the addition of these two variables in a regression equation contributed only negligibly to the multiple correlation coefficient.

In general, analyses of the subtests of the PMA test and Raven matrices yielded the same conclusions. Grade was the only significant contributor to regression equations involving five of the seven subtests for which data was available from all subjects.

Grade was *not* a significant predictor in only two cases. Sex ($r = .55$) was the only significant predictor of performance on the verbal analogies subtest (no. 1), and age ($r = .33$) was the only contributor to the reasoning subtest (no. 4). In only one case (the recognition of figures subtest no. 8a) did the use of multiple correlation techniques improve the fit of the regression equation. While grade ($r = .42$) was the largest contributor of the multiple correlation of .63, inclusion of both sex and age improved the final result.

Discussion.—These results are in general agreement with the data from the individual experiments reported in earlier sections of this *Monograph*. Years of educational experience is by far the most pervasively effective demographic variable related to performance on complex tests.

The two exceptions to this generalization in the case of PMA subtest performance are difficult to interpret. On the surface, the reasoning subtest is quite similar to the verbal logical problems study reported above. Yet, in the case of the problems from the PMA test, age, not grade, was the primary predictor of performance. Only further research could tell us whether this result arises from differences in the way the problems were constructed, the range of age and educational experience sampled, or some other factor.

The emergence of sex as a major determiner of the verbal analogies performance has only one counterpart in our studies. Sex emerged as a minor, but reliable, predictor of performance on the verbal similarities problems, which we used as a measure of IQ in the survey to be described in the next section. The two tasks appear quite similar. The reasons for sex as a significant predictor are at this point completely open to speculation.

V. DETERMINANTS OF EDUCATIONAL ACHIEVEMENT

Taken at face value, the evidence from our various experiments seems to require the conclusion that formal schooling does, indeed, improve test performance on a variety of cognitive tasks.

In Chapter VI of this *Monograph*, we will discuss various explanations and interpretations of this improvement, but first we need to examine the fundamental question: Is it schooling that is responsible for the difference in performance associated with grade in our experiments?

The major alternative hypothesis is that schooling operates only as a screening device: as we increase grade we are working with a smaller and smaller sample of the population which is more and more highly selected for intelligence, the ability that produces school success.

We could argue that the experimental data themselves render such an interpretation very implausible because they would pose irresolvable paradoxes and very unlikely conclusions. We would have to conclude, for example, that uneducated Yucatecans fail to develop memory capacities beyond the level associated in the United States with that of a normal 8-year-old child. We would have to conclude that people who live in small towns are less intelligent than people who live in larger towns, or that people who live in towns where no secondary school has been built are less intelligent than those who live in towns with secondary schools.

Whatever the plausibility of such conclusions, the problem of selectivity in our various subject populations must concern us if we want to assign causal status to the effects we have associated with formal schooling. Ideally (in terms of experimental design), we would like to be able to show that there were no selective factors associated with grade, but we know that this is not the case, even judging from the demographic information we presented earlier. Realistically, then, we want to determine the selective factors at work in Yucatan to find out if they undermine our conclusions about the effects of formal education.

THE DEMOGRAPHIC SURVEY

A sample of 393 adults between the ages of 16 and 85 years of age were administered a questionnaire to provide systematic data on the demographic variables associated with varying amounts of schooling. None of these people was enrolled in school at the time of the interview. The entire questionnaire is reproduced here as the Appendix. Most of the early questions concern the interviewees' family backgrounds and education experiences (including occupation and educational status of their parents, availability of education in their hometown, regularity of their attendance in school), while later questions deal more with their attitudes toward education for themselves and for modern Mexico. Mixed into later questions is an adaptation of a vocabulary subscale of the WAIS.

The towns where the interviewing was conducted and the number of people interviewed in each town are listed in figure 1 (p. 5). We attempted to interview at least 5% of the adult population in each town using a modified strata-sampling procedure. The interviewer obtained data from four people in the age range of 16–30 years, two people in the range of 40–56, and two people in the range above 56 years of age for every 10 persons sampled. The interviewer also made sure that women were adequately, but not necessarily equally, represented. Naturally, this sample included only people who would cooperate with the interviewers. Noncooperation was minimal but relatively more common among women than men. Each interviewee was paid 20 pesos ($1.60) for his or her participation (a figure that for a majority of our subjects represented better than 50% of a day's wages). Sampling in the larger towns included roughly equal numbers of interviewees from each barrio.

Finally, it should be stated that the objective of this survey was not to obtain a sample representative of the population of Yucatan as a whole (although we feel that this objective has been fairly well accomplished). Rather, we were concerned primarily to obtain a sample that was representative of the population from which our experimental subjects were drawn.

Table 13 presents the average scores for several demographic variables including those found to influence the level of educational attainment and responses to the WAIS subtest.

A few facts are clear even from the summary results represented in table 13. Maya is the first language of about three-fourths of the sample, but only a half currently speak Mayan at home. Interviewees have more education than their parents, but the average is low and less than that available in their hometowns.

Of more concern than these data on the average response to various questionnaire items was to determine how these variables related to each other. A rough answer to this question is provided in table 14, which con-

TABLE 13

Characteristics of Survey Sample

Variable	Average Value	Range
% married	82	0–1
N siblings	3.9	0–16
Age	37.6	16–85
% Mayan as first language	78	0–1
% Mayan as current home language	46	0–1
Father's occupation[a]	2.1	1–6
Father's highest grade	1.5	0–16
Mother's highest grade	1.0	0–14
Interviewee's highest grade	2.6	0–17
Highest grade in hometown	5.0	0–16
Subject's occupation[a]	2.3	1–6
% male	61	0–1
WAIS score	8.5	0–26
Modernity score	5.4	0–17

[a] Occupations were scaled on a six-point scale as follows: 1 = self-employed farmer, 2 = unskilled nonfarm laborer, 3 = semiskilled craft, 4/5 = skilled craft/government worker (two levels distinguished), 6 = commercial/professional.

tains the first-order correlations among the variables listed in table 13. All correlations represent approximately a .01 significance level.

As the large number of entries in the table readily indicates, number of years of schooling ("highest grade") is positively related to parental education, father's occupation, and the availability of schooling in the town. It is negatively related to age, the use of Maya in the home, and being married. These and other conclusions, while interesting in themselves, tell us only that many background factors are associated with educational attainment. These background factors, in turn, are related to each other, rendering this first-order correlation matrix of limited value in specifying the mix of factors that determine how many years of education members of our sample obtained. Nevertheless, it is instructive to examine the various correlations for confirmation of the general statements about the relationships between demographic factors which were made in earlier chapters of this *Monograph*.

For example, our contention that Mayans are a less privileged part of the population is confirmed by the many significant negative correlations between speaking Maya in the home and factors related to income (occupation, schooling, living in a town with a high level of schooling available, etc.). Older people experienced less education (a fact reflected in table 13 as well), but from the correlational data we can see that the negative correlation between highest grade and age is not as large as that for age and highest grade in town (hereafter "TOWN.GR"). This result suggests that the greater availability of schools is a major factor in the age-related increase in

TABLE 14

CORRELATIONS AMONG MAJOR SURVEY VARIABLES

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Married			.33						−.21	−.16			
2. *N* siblings			−.34										
3. Age					.15			−.16	−.28	−.41			
4. Mayan first language					.42				−.45	−.26	−.24		−.23
5. Mayan home language						−.43	−.32	−.31	−.46	−.34	−.33		−.28
6. Father's occupation									.57	.35	.36		.35
7. Father's highest grade									.42	.34	.26		.21
8. Mother's highest grade									.50	.30	.30		.21
9. Interviewee's highest grade										.55	.50	.16	.36
10. Highest grade in hometown											.36		.22
11. Subject's occupation												.44	.32
12. Male													.22
13. WAIS													

education. Men attain higher grades than women and much higher occupational status; they also score higher than women on the WAIS subtest.

We could continue this list, but the many intercorrelations between background factors makes their interpretation virtually impossible. Consequently, we decided to analyze these data using multiple-regression techniques organized around an explicit model of the factors controlling highest grade.

Figure 9 represents the processes that we initially hypothesized as determiners of highest grade. We will speak of factors which are linked to highest grade by a single arrow as *direct* determiners, and those which are linked by highest grade via another factor as *indirect* determiners of highest grade.

Our model assumes five direct determiners and seven indirect determiners of highest grade. We will discuss these in turn, elaborating on each as we go.

1. Town grade (TOWN.GR). The amount of education available in different towns, at the time when the informant was growing up, varies considerably. We assume that this variable will act like a barrier and that it should be positively related to highest grade.

2. Mayan first language (MAYA.1STLANG). We assume that those people who grew up in homes where only Maya was spoken will suffer a disadvantage in a school system where all the instruction is in Spanish.

3. Financial constraints. Education has an economic cost for families, both directly because of expenses for books and clothing, and indirectly through the opportunity cost of giving up the labor services of the child. We do not have a direct measure of the financial constraint variable, but we can estimate one of its principal components (family income) and measure another (family size).

4. Value of education. This variable designates the complex of values

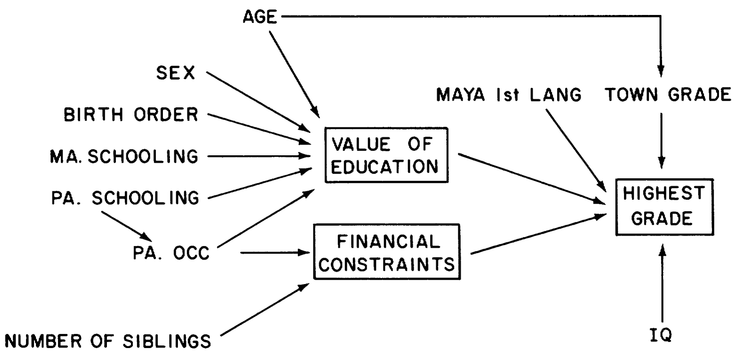

FIGURE 9.—A schematic model of the antecedents of last grade (see text for explanation).

that a family associates with education: their notion of how important it is, how much a child ought to receive, and how much sacrifice they are willing to make to provide that education. We do not have a direct measure of this variable, but we can measure six of the factors which determine its size: parental education and occupation (which ought to increase the value attached to education), and sex, age, and birth order of subject.

5. IQ. Our measure of IQ is the WAIS subscale score. Since this score has not been standardized for the populations in question and because it may be as much a consequence of education as its cause, inclusion as a determiner is highly suspect. We include it for heuristic value, recognizing the limitations it carries with it. We comment further on this below.

We turn now to the seven indirect determinants of highest grade, which are hypothesized to work as financial or values factors.

1. Family income cannot be measured directly in this sample, but we do have a proxy for it—father's occupation (PA.OCC). This variable is coded on a six-point scale, with the value of 1 being used for subsistence farmers and the value 6 for professionals. The higher the value of the index, the higher the expected income. A separate interview study in which 20 informants were asked the social prestige of 32 occupations and the daily salary associated with each produced a correlation of .77 between rated prestige and salary. Father's occupation is assumed to affect the desired level of schooling in addition to increasing financial resources because higher occupation codes are likely to involve more contact with the outside world and greater perception of the uses of education.

2. Number of siblings (NO.SIBS) (family size) ought to influence financial constraints, and it seems reasonable to assume that it will interact with family income; the higher one's income, the less of a constraint it should be. To represent this factor we created $VARIABLE-1, which is father's occupation divided by number of siblings, that is, this approximates income per child.

3 and 4. Father's schooling (PA.SCH) and mother's schooling (MA. SCH) are assumed to influence educational values in a positive direction. The more educated the parent, the higher the educational goals for the child. Father's schooling almost certainly affects his occupation and thus his income, whereas mother's education is less likely to have such an indirect effect on finances because few mothers work outside the home.

5. Sex is coded as 0 for females, 1 for males, and is assumed to affect desired amount of education positively; higher goals are set for male children than female children, who marry earlier and whose labor is required at home.

6. Age (subject's current age) is meant to serve as a proxy for several factors. We assume that general belief in the value of education has increased over time, in addition to the increased availability of schools. Hence, age

indicates how long ago a family's decision about educating a child would have been made. The age variable should be negative.

7. Birth order (B.ORDER) should affect the desired amount of education for the child, with the highest goals for the firstborn children. This variable is coded 1, 2, 3, depending on whether the subject was born in the first, middle, or the last third among his siblings. We assume that birth order will also interact with family finances: the more children one has, the heavier the burden of late children when it comes time to pay school fees. Various versions of an interactive variable which takes into account income, family size, and birth order were tried; we settled on $VARIABLE-2, which is $VARIABLE-1 divided by birth order.

Although figure 9 implies a general form for the regression equation, we needed to make decisions regarding the exact form of two of the financial variables containing father's occupation, and the form of the age variable. So we began with an equation which included all the variables from figure 9 and did some experimentation with these three variables to find the form which gave the best fit to the data. The age variable is meant to represent a time trend, and it is not obvious whether the time trend is linear or not, so we experimented with using age alone, age^2 alone, as well as age and age^2 together. A simple, linear age trend worked the best. We soon arrived at the two financial variables involving father's occupation: the interaction between PA.OCC and NO.SIBS ($VARIABLE-1), and the interaction between PA.OCC, NO.SIBS, and B.ORDER ($VARIABLE-2). In each case we experimented with a variety of nonlinear transformations of the variables until we found the form which gave the best fit as described above.

The equation obtained using these variables produced an $R^2 = .562$ (corrected for *df*). It was:

$$\begin{aligned}
\text{GRADE} = 1.34 &+ \underset{(6.8)}{.256\ \text{TOWN.GR}} - \underset{(4.4)}{\text{MAYA.1STLANG}} \\
&+ \underset{(4.4)}{.249\ \text{MA.SCH}} + \underset{(2.9)}{.549\ \text{SEX}} - \underset{(3.0)}{.0200\ \text{AGE}} \\
&- \underset{(3.6)}{.341\ \$\text{VARIABLE-1}} - \underset{(4.2)}{.385\ \$\text{VARIABLE-2}} \\
&+ \underset{(3.2)}{.776\ \text{PA.OCC3/4}} + \underset{(3.0)}{1.52\ \text{PA.OCC5/6}} \\
&+ \underset{(3.3)}{.0558\ \text{IQ}}
\end{aligned}$$

The number in parentheses below each regression coefficient is the t ratio, b/s_b. The equation was fitted by ordinary least squares, with no constraints on the coefficients. Examining this equation we note that all of the regression coefficients have the theoretically expected signs, all have plausible magnitudes, and all of them are significant at well beyond the 95% level. The R^2 (multiple $r = .75$) says that we have explained 56.2% of the variance in the data, which is an extraordinarily good fit for data of this kind (i.e., cross-sections of individuals where the dependent variable has a limited range and is truncated on one end).

We will discuss each of the regression coefficients in turn. The regression coefficient of TOWN.GR is positive, as expected, and each additional year of schooling available in the town will produce an expected increase of .256 years for the average individual, all else held constant. (Although .256 is the value of the regression coefficient, it should be understood that we are not estimating a real world effect with three-digit accuracy. Regression coefficients are subject to sampling variation and may be sensitive to model specification as well. Our experience with this equation on a variety of subsamples of the data, and with some experimentation in the specification, shows that the regression coefficients are in fact relatively stable and tend to maintain relative sizes when they do vary. Obviously this would not be true for β coefficients, but we are not using them at this point.)

MAYA.1STLANG was coded as 1 for those who grew up in households where only Mayan was spoken and coded as 0 for others. Its regression coefficient implies that Maya-speaking homes are associated with an expected handicap of 1.1 years of school, all else held constant. MA.SCH is positive, as expected; each additional year of mother's education raised the expected amount of the subject's education by .25 years, all else held constant. (Father's education would not enter the regression in a significant manner, $t = .4$, and so it was dropped from the equation. Apparently, once the effects of father's occupation are held constant, there is no separate variance left for PA.SCH to explain.) SEX is coded 1 for men and 0 for women. Its coefficient is positive, as expected; holding all other factors constant, we would expect that women would receive .55 years less schooling than men. AGE is negative, as expected, showing a slight positive time trend in people's value of education. (Given the range of ages in our sample, this trend could account for a maximum difference of 1.4 years of education.) The next two variables, $VARIABLE-1 and $VARIABLE-2, both represent interactions of family size and the father's education and are easiest to interpret if looked at jointly. The net combined effect is a positive .13 years of schooling at the average value of the explanatory variables; and there is a possible effect of 13 years of educational difference across the extreme values of the variables. The next two variables are different ways of coding the level of father's occupation, and they can best be interpreted with respect to someone whose father was in occupation 1 or 2 (a farmer or a domestic worker). The first (PA.OCC3/4) says that someone whose father is in occupation 3 or 4 (tailor, shoemaker) would expect to receive .78 more years of education than someone whose father was in occupation 1 or 2, holding constant all other variables. The next coefficient (PA.OCC5/6) says that someone whose father was in occupation 5 or 6 (doctor, teacher) would expect to receive 1.5 more years of education than someone whose father was in occupation 1 or 2, all else held constant.

Before going on to discuss the general implications of the regression equation, we want to look more closely at some of its components.

THE IMPORTANCE OF IQ

Our various tests and experiments have shown a strong association between cognitive performance and years of education. We have been interpreting this as a causal relationship, though we pointed out that one alternative hypothesis might also account for the data: the notion of IQ screening, that is, further years of education operate progressively to screen out the less-able members of the population. By this alternative explanation, the underlying cause of the high cognitive scores is the school system's selection effect rather than its educational effect, or a demographic effect not related to initial ability.

The regression analysis provides a possible way of distinguishing between these hypotheses. The IQ-screening hypothesis predicts that the regression coefficient for the IQ variable ought to be positive, significant, and important, while the other hypothesis makes no particular predictions.

There are, however, two obvious qualifications which must be placed on such a regression analysis: (*a*) our IQ measure is far from perfect since it has not been standardized on the population in question and is only a subscale in any event, and (*b*) we expect higher scores for educated subjects because education is training for the kinds of thinking which the IQ test measures.

Although our IQ measure is clearly a proxy, it is by no means an insupportable choice. In an extensive study of American and Mexican children, Holtzman, Diaz-Guerrero, and Swartz (1975, pp. 111 ff.) demonstrated that the vocabulary subscale of the Wechsler Adult Intelligence Scale (WISC) entered into roughly the same factor structure in both populations and that the vocabulary subscale correlated significantly with such different scales as block design and comprehension. These results cannot, of course, be applied uncritically to our samples since we worked in a different part of Mexico with different indigenous Indian populations. To the extent that our vocabulary test is an inappropriate proxy for IQ, it will tend to bias the regression coefficient toward insignificance. And under the extreme case where there was no relationship between vocabulary and IQ, the regression coefficient would not show any association at all and would be insignificant.

The second effect, the spurious correlation between *present* IQ and *past* education, will work in the opposite direction. Even if there is no IQ-screening effect whatsoever, this kind of testing bias will produce an association between measured IQ and education and hence operate to bias the regression coefficient toward being spuriously large and significant. Thus either an IQ-screening effect, or an education-for-taking-IQ-tests effect, would tend to produce a regression coefficient which is large and significant. If we do obtain such a coefficient, we cannot easily know the real cause. But if the resulting regression coefficient is insignificant, or relatively unimportant, then we need not worry about differentiating the two effects; we will know that IQ screening is not a major determinant of our cognitive testing results.

The basic regression equation above already includes our IQ proxy variable, and its regression coefficient is statistically significant; but, is it important? Unfortunately there is no agreed upon, methodologically correct procedure for measuring importance. Four possible measures—contribution to R^2, relative β weights, zero-order correlation, and possible range of effect—have been used in the literature in various ways. The fourth measure is the most robust, the first three being more sensitive to sampling effects, but all four have something to say about some aspect of the question, and so we utilize all four.

Contribution to R^2.—If we remove IQ from the equation, the R^2 drops by .011, which is clearly not much of a change. This test is, of course, highly sensitive to variations in sampling and especially to the mix of other variables that are present or absent in the equation. In general, there is some bias toward understating the effect, since mutual correlation between explanatory variables will act to minimize the change in R^2.

Relative β *weights*.—Although nothing can be inferred from the relative sizes of the ordinary regression coefficients, it is reasonable to compare the relative size of the β weights, the standardized regression coefficients. Again, this test is subject to sampling variations: if there is little variation in some explanatory variable in a particular sample, then we will tend to find a low β no matter how important it is. In this case, we want to compare the β for IQ to the β for all of the other variables combined. To do so we define a new variable, W, which is the weighted sum of all the other variables in the equation, where the ordinary regression coefficients are used as the combining weights. Thus W will exactly measure the combined effect of all the other variables. We then regress grade on IQ and W and observe the relative β's. The β coefficient of IQ using this procedure is .12, while the β coefficient of W is .71; that is, a 1 standard-deviation change in the rest of the variables will produce an effect on grade which is six times bigger than the effect of a 1 standard-deviation change in IQ.

Zero-order correlations.—Unless there are unusually effective suppressor variables operating, the zero-order correlation will generally overestimate the effect of a variable because the zero-order correlation includes all the indirect correlations between the two variables which operate through other factors. The zero-order correlation between IQ and years of school completed is .36, and if we square this we can calculate that IQ could account for only 13% of the variation in schooling, holding *nothing* else constant.

Range of effect.—The ordinary-regression coefficient is relatively robust across samples since it is a slope estimate. We know that the possible score range on our IQ proxy is 26 points. If we multiply this range by the computed slope estimate for IQ, we calculate a maximum possible effect of only 1.45 years of education. That is, the predicted variation in schooling attributable to the maximum possible variation in IQ is only 1.45 years. (Note that using an IQ scale with a greater range would not change this result, since the regression coefficient would automatically adjust downward to keep the total effect constant.)

Thus on all four of our possible measures of importance, the IQ proxy measure is shown to be relatively nonimportant. We also tried some alternative regression analyses to examine this same point.

It is possible that the observed nonimportance of IQ is a result of the range of schooling in the sample: there are many people who have zero years of schooling because of lack of opportunity, not lack of intelligence; it is also possible that IQ is only a significant factor at higher levels of education. To clarify either of these effects we deleted all subjects who had less than 3 years of education, leaving a grade range of 3–17 years and a sample size of 180. We then ran our basic equation on this screened-down sample to see if the measured effects of IQ would be significantly increased. They were not: the marginal contribution to R^2 increased slightly to .016, β rose slightly to .14, but the possible range of effect fell to 1.23 years, and the t ratio fell greatly from 3.3 to 2.4.

To make one final test of the importance of IQ, we screened the sample down to the 116 subjects who came from towns that had exactly 6 years of education available. We did this to hold constant the wide variety of complicating effects which might be associated with differential schooling opportunity. However, this subsample made IQ look even less important: the zero-order correlation between IQ and grade fell to only .27, and the computed regression coefficient of IQ was not statistically significant when other factors were held constant.

In conclusion, we have performed a variety of tests and run a number of regressions to examine the possible importance of IQ in the determination of grade. All of the tests and equations point in the direction of nonimportance. This analysis is subject to two qualifications: our IQ measure is only a proxy for IQ, and to the extent that it is a poor proxy there will be some bias in the direction of showing nonimportance. On the other hand, since there is good reason to believe in a spurious causality—increasing education produces increased ability to respond in the manner which IQ tests measure—

there would be some countervailing biases in the direction of overstating importance. Despite the problems in our proxy measure for IQ, it would need to have almost zero relationship to a "true" IQ measure in order to explain away our finding that IQ is nonimportant. On balance, then, after allowing for the possible imperfections of our measure, we still strongly conclude that the relationship between IQ and educational achievement is small, within the range of grades in our sample.

TOWN GRADE

We had originally put this variable into the equation under the assumption that it represents a kind of barrier; it sets a limit to the amount of schooling which a subject can obtain at a reasonable cost since migration or boarding are expensive. As such, we expected that it would be the dominant variable in our equation and that it would explain almost all of the observed variance. However, neither its regression coefficient, .26, nor its β coefficient, .29, are as high as we would have expected. Looking back to table 13, we see that highest grade is considerably lower than town grade.

Figure 10 shows the distribution of highest grade plotted against town grade. The diagonal line marks the point at which the number of grades completed exactly matches the number of grades available. From the graph it is apparent that very few people in our sample left their towns to achieve

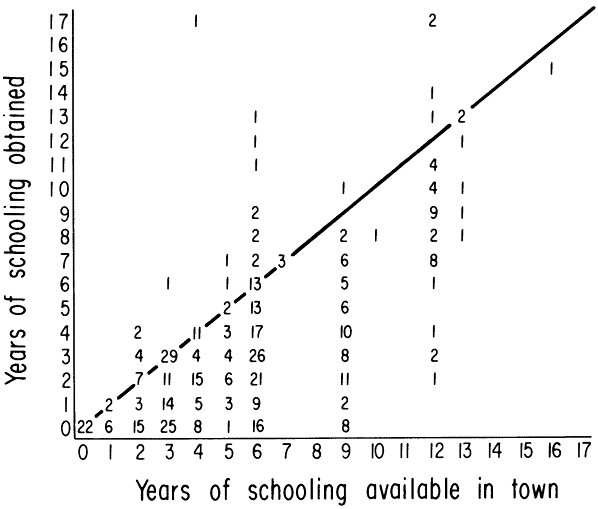

FIGURE 10.—Last grade of school completed plotted as a function of the amount available in hometown. The discontinuities on the abscissa reflect the fact that schools are organized in units of three grades. The presence of entries at other grades reflects both the unavailability of teachers and errors in subjects' reports.

more education; this result strongly supports the intuitively appealing idea that town grade is a barrier to further education.

However, the graph also clearly shows that most people failed to complete the grades available to them, often falling one or two years short of "graduation."

This observation led us to surmise that, in addition to acting as a barrier, town grade may also function as a means of determining people's educational aspirations. Many of the towns we sampled are somewhat isolated from the modern world, and under such conditions the local facilities might act to set local norms.

The questionnaire contained information about *current* educational values (e.g., "How much education do you believe someone ought to have?"), and so we decided to use this question as a proxy for past educational values. To make this procedure more reasonable, we excluded everyone above the age of 25, so that a closer relationship between the measured variable (current values) and the variable of interest (value of education at the time the subject was in school) would be plausible. Then we used current value as the dependent variable and regressed it on TOWN.GR plus other background factors. We found that (1) current desired level of education is equal to about 70% of TOWN.GR, and (2) TOWN.GR is the most important factor in explaining the variance in current educational aspiration. We take these two findings as support, though certainly not proof, for the "aspiration" function of TOWN.GR. That is, years of schooling available is an objective, easily seen marker in a situation where people would otherwise have little information with which to set goals for their children.

LANGUAGE

The effects of the language variable (MAYA.1STLANG) are worth exploring in some detail because of their general relevance to bicultural and bilingual educational environments.

It was not uncommon for us to encounter village schools where the teacher spoke no Mayan and the students spoke little, if any, Spanish. Hence, we assumed that growing up in a Maya-speaking home would be an educational disadvantage that would retard the acquisition of basic skills and promote school leaving. We were very uncertain, however, if the negative consequences of speaking Maya were distinguishable from the negative effects of low income, family size, and the other factors associated with lower educational achievement.

To clarify the role of MAYA.1STLANG, we ran a series of regression equations to determine at what stage of educational experience Maya speakers encountered difficulties severe enough to discourage further attendance. Two rival hypotheses seemed plausible. If Maya speakers could not adjust to Spanish quickly, MAYA.1STLANG would operate heavily at the lower grades where children would be expected to learn to read. On the other hand, teachers might be reluctant to discourage the young newcomers in order to give them a chance to master the rudiments necessary for later work; only when the more advanced work, which presupposed reading knowledge of Spanish, became essential would Maya speakers be heavily penalized. By this latter hypothesis, MAYA.1STLANG would become operative at the third or fourth grade.

To determine when MAYA.1STLANG began to exert its negative influence on GRADE, we split our sample at various grade levels to determine when the coefficient of MAYA.1STLANG becomes significant.

We obtained the following results: 0–1 years of education, MAYA.1STLANG not significant; 0–2 years of education, MAYA.1STLANG barely significant ($p = .05$); 2–17 years of education, MAYA.1STLANG highly significant; 4–17 years of education, MAYA.1STLANG highly significant.

Whether as a result of curriculum demands or because MAYA.1STLANG operates as a proxy for more traditional values which do not emphasize formal education, the negative educational effects of speaking Maya as a first language appears only after the completion of 2 years of education.

CONCLUSION

We would be foolhardy on the basis of these results to conclude that our schooled and unschooled populations are random samples of the Yucatecan population. They clearly are not; moreover, there are different mixes of demographic variables associated with different levels of schooling.

What we can argue against on the basis of these data is the view that in working with people at different educational levels we are doing no more than selecting people according to their predetermined intelligence.

By far the most severe constraint on the amount of schooling obtained is the amount available. Next are a set of economic factors which are so severe at the lower-income levels as to make one marvel at parents' faith in making the sacrifices necessary to provide their children with an education.

While we cannot rule out the operation of factors which select for intelligence and are only indirectly related to school attendance (Mayan farmers are less likely to live in towns with a good deal of schooling available), we would have to posit a population for which the statistically normal levels of intellectual functioning corresponded to that of a 6-year-old middle-class American child in order to bring our conclusions in line with the experimental data presented in earlier chapters. That possibility seems sufficiently remote to shift the burden of its proof to those who would support it.

We are content to rest on the evidence that the level of one's education in Yucatan is in large measure determined by economic and cultural factors over which the individual has relatively little control, and that intellectual test-performance differences among our populations are best attributed to differences in education and work experience. However, we realize that, insofar as factors like mother's education or father's occupation are acting as surrogates for IQ, and insofar as our test of IQ is unreliable, the evidence is equivocal. In our opinion, the inferential dilemmas these uncertainties pose for us cannot be resolved until and unless extensive longitudinal studies such as those begun by Irwin, Engle, Klein, and Yarbrough (1978) have been completed.

VI. GENERAL DISCUSSION

A BRIEF SUMMARY

The preceding sections have been heavily packed with data. We now want to turn to issues of interpretation. First, however, we present the following summary in order to help the reader keep the salient experimental results in mind for purposes of discussion.

1. When categorizing a set of 20 pictures of common objects all population groups sort taxonomically, but GRADE seems to produce tighter, more organized sortings: uneducated adults mix functional and taxonomic organization of the stimuli.

2. When asked to group artificially constructed stimuli (triangles-squares, red-blue, etc.), subject's performance is very closely associated with grade; age appears irrelevant for the age range studied. Recourse to "natural" stimuli (red and yellow corn kernels) for which these distinctions are commonly made, but are without pragmatic significance in the task we presented, leaves the results unchanged.

3. If subjects are asked to match three objects with a target object on the basis of their unspecified similarity ("choose the object that belongs with this hat"), performance is a function of age (and the nature of the distractor items). Correct choices were virtually always justified in terms of the functional relations between items.

4. If two semantically related items were verbally presented, along with a single distractor, and subjects were required to select the two that belong together or the one that is different, both increased age and grade were associated with improved performance, but the magnitude of their effects depended upon the nature of the distractor and the aspect of performance one examines: differences in response choice were more strongly associated with grade than age, but both age and grade were associated with increases in category-based verbal justification of subjects' choices. Subjects in all groups justified noncategorical choices by naming functions common to the two items chosen in a majority of cases. Analogous results were replicated using procedures derived from the WAIS.

5. In free-association studies, all groups tend to give noun responses to noun stimuli; only high levels of education differentiated groups with respect to paradigmatic responses to adjectives and verbs. Within-semantic-category noun responses to noun stimuli were at a substantial level only for the most educated groups in these studies; all groups gave many noun responses that were functionally related to the stimulus words.

6. Free-recall performance varied as a function of grade. Especially marked was the fact that high levels of semantically clustered recall were observed only for secondary school students. In paired-associate recall, the grade effect was most marked when stimuli and responses were randomly paired with respect to semantic class.

7. Short-term recall of location varied as a function of grade. Fine-grain analysis indicating differential amounts of primacy implicated active rehearsal as the behavior which differentiated more and less education subjects.

8. Responses to verbal logical problems varied as a function of grade, and perhaps of exposure to modern commerce. The more sophisticated groups were distinguished by their strong tendency to treat the problems as a logical puzzle to be solved on their own terms. Less-educated groups were more likely to respond in terms of their knowledge of the world around them.

9. Total scores on Thurstone's PMA test varied as a function of grade.

Whatever the shortcomings of individual studies (and there are many), the evidence from the entire series of experiments, combined with our data on the factors limiting years of schooling, make it appear reasonable to conclude that differences in educational experience make a substantial difference in the intellectual skills that subjects bring to bear in our individual tasks. Assuming this to be true, two questions arise: (*a*) How can we best characterize the nature of the behavioral differences we observed between schooled and unschooled subjects, keeping in mind that these differences were not uniform across tasks. (*b*) Does schooling promote the development of intellectual skills in general, or are we simply recording the outcome of differential practice with tasks to which we should attach no general significance?

CHARACTERIZING BEHAVIORAL DIFFERENCES

Does Education Influence Stimulus Familiarity and Lexical Organization?

The current literature on education and cognitive development suggests two major lines for the interpretation of these results within a general framework which assumes that test results reflect general cognitive skills: one emphasizes difference in subjects' knowledge about the stimulus materials; the other emphasizes differences in the way stimulus information is processed.

Hall (1972) obtained results very similar to those reported here in a study conducted among Colombian children of different ages and levels of educational attainment; he interpreted his data as a reflection of differences in the attributes of stimuli to which his two populations attended. In a paired-associate study similar to our "between-category" paired-associate condition, recognition memory, and a free-association study, he observed substantial differences associated with education. Like us, he found that his more educated subjects gave superordinate responses in free associating to nouns and more paradigmatic responses to adjectives. He also observed that his more educated subjects were likely to err in their recognition memory by falsely asserting that they had previously seen items that were only likely associates of items that had really been presented. The only task for which education-related differences were small required the subjects to learn a set of verbal discriminations.

Hall concluded that the relative lack of differences in the verbal discrimination task (where research has shown frequency of exposure to be the critical variable), coupled with large differences in tasks where sensitivity to the "verbal associative and acoustic attributes" of stimuli are important, implicates more effective use of these latter stimulus features by the educated subjects as the source of his group differences. Something about formal education induces the educated subjects to attend to stimulus attributes which are more effective instruments of remembering (Hall mentions "the emphasis during the early school years on conceptual and associative relationships among words" as a likely causal variable). Fjellman (1971) also emphasized the importance of stimulus familiarity, observing superior performance among noneducated subjects in classifying tasks when the educated subjects were less familiar with the objects to be classified. We, too, have been impressed with the possibility that differences in the content and structure of people's knowledge about our experimental materials and tasks might underlie the performance differences we have reported. It is for this reason that we took pains to use culturally familiar materials so that differential knowledge of test content would be minimal in our categorizing and memory studies.

The results of these efforts, as reflected in our data summary, do not build a very strong case for interpreting the pattern of education-related performance differences in terms of differential familiarity with the task stimuli.

Consider the data from the free-association studies, where our results and Hall's are in close agreement. A central index in these studies was the probability that people would respond to each stimulus with a response from the same grammatical class (termed paradigmatic responding, e.g., red-green or house-tree). When nouns were the stimuli, all groups tended to give noun (paradigmatic) responses. But for adjectives, in Hall's study, and adjectives and verbs, in our work, the more highly educated subjects responded paradigmatically more often than uneducated subjects. What theory of differential stimulus encoding can encompass differences in responding conditional on grammatical word class?

The results of our study in which people had to form categories of geometric figures or pieces of maize must also appear puzzling from the perspective of cultural familiarity. The color and size of corn kernels are both salient attributes of these objects for Mayan farmers, attributes which they readily use in describing individual stimuli, just as students are used to describing geometric figures. But farmers were no more successful at forming groups with maize than they were with geometric figures, nor did the students deal more successfully with the geometric stimuli.

It is findings such as these that have made us skeptical that differences in lexical structure or stimulus familiarity is a general explanation for the education-related differences observed in our studies (see, also, Fjellman [1971], who found education-associated differences not explainable by her indices of familiarity).

It's What You Do with What You Have That Counts

At the risk of drastically oversimplifying a very complicated set of relations, we would like to suggest that the major results of our research are captured by the notion that more highly educated subjects more readily engage in intellectual activities which are not rigidly predetermined by the structure of the task and which promote efficient performance. It is not differences in the information about the stimuli per se but differences in what people do with commonly available information that is critical. If we are correct, differences between educated and noneducated subjects will depend on the extent to which the task permits or requires such activities and the difficulty of the required behaviors. If the task itself organizes subjects' responses so that it can be solved without recourse to special intellectual work, or if the task's demands are beyond the reach of the educated subjects, little or no education-related differences in performance are to be expected.

In various guises, this line of thought can be found in a great variety of current cognitive theorizing (see Brown 1975a, 1975b; Flavell & Wellman 1977). As a rule, it has been used to explain age-related differences in the performance of children, but it has been applied in comparisons of retarded and normal children (Brown 1975a) and children from different social class or ethnic groups (Jensen 1969). Earlier applications to education-related differences can be found in Cole et al. (1971), Cole and Scribner (1974), Scribner and Cole (1973), while a more sophisticated recent discussion is presented by Brown (1977).

The nine-item summary of results with which we began this section is a useful reference point against which to evaluate the usefulness of our loosely formulated "cognitive-processing" interpretation.

The following conditions failed to produce marked education-related differences in performance: (1) the sorting studies where subjects were asked to group items from large, recognized semantic categories on the basis of their similar meaning, or when functional sorting produced performance that is indistinguishable from semantic sorting; (2) when a paired-associate list was constructed using (presumably) highly associated items from the same semantic category; (3) in a short-term memory study, when the location presented just prior to recall was the probed-for item.

All these tasks (or parts of tasks) share the characteristic that the basis of solution is made obvious by the lack of competing alternatives, or that little processing is required to produce a response (as in the short-term recall study where the last-presented item is often said to reside in "echoic" memory).

Marked education-related differences in categorizing appeared in several different contexts: they were apparent in categorization tasks containing several competing sorting principles, especially where functional sorting criteria were absent, or when semantic and functional classification were pitted against each other.

Education-related differences in memory tasks occurred in close association with those aspects of performance widely accepted as evidence of "strategic" activity (Brown 1975a, 1975b). In the free-recall task, the effect of education was most strongly manifested in reordering of the to-be-recalled list to conform with its latent taxonomic categories. In the paired-associate study, the superiority of the educated subjects was greatest for the more *difficult* list, consistent with Rohwer's (1973, pp. 1–57) repeated observation that group differences appear under circumstances where subjects must "elaborate" the list items to facilitate their association. In the short-term location-recall study, education effects occurred for early and middle items in the list—that part of the task where spontaneous, cumulative rehearsal has been demonstrated to be an essential accompaniment of enhanced performance in older children (Hagen & Hale 1973). All of these results are consistent with the generalization that the more highly educated subjects are engaging in deliberate remembering activities which successfully augment their performance.

In the verbal problem-solving task, the evidence clearly points to an education-related difference which hinges on the evidence subjects deem necessary to make a judgment; a few years of educational experience appears sufficient to induce people to treat these questions as a problem in reasoning from the given premises to a conclusion, while uneducated subjects seek empirical support for their evaluations.

The PMA test is an amalgam of tasks like those we have just discussed, and the effects of education are very clear.

Does Education Promote Cognitive Development?

So far in this *Monograph* we have been extremely careful in the language we have used to characterize the performance differences between educated and noneducated subjects. Even in the preceding section where we urged differences in cognitive operations as the locus of education effects, not differences in stimulus familiarity, stimulus attributes, or lexical organization (all versions of the position that the content of the tasks are not equivalent across groups), we were circumspect in our claims for the generality of the results we were reporting. Here we want to address directly the issue of generality. Are we observing and reporting phenomena that are characteristic of the thinking of educated and noneducated Yucatecans? Or are we dealing with a circumscribed set of tasks which have little significance outside of experiments (and possibly schools)?

For most researchers who have considered the problem of the effects of education on cognitive development, the conclusion to be reached from this kind of research is clear. Hall (1972), for example, concluded that formal education is responsible for producing differences in what he terms "verbal development" (p. 287). Wagner (1974), in his summary of the short-term memory experiment described earlier, concluded that "higher mnemonic strategies in memory may do more than 'lag' by several years [without education]—the present data indicate that without formal schooling, such skills may not develop at all" (p. 395). Brown (1977) generalizes this conclusion in a way that is compatible with the results of our research: "Even this brief consideration of cognitive development in cultural perspective suggests that much of what we regard as the 'normal' course of development is, if not actually an outcome of formal schooling, at least greatly influenced by the process" (p. 13).

Our doubts about this kind of conclusion come from two sources, one largely empirical, the other speculative (theoretical would be too respectable a word).

Empirical Inconsistencies

On the empirical side, we have to consider apparent contradictions to our generalization based on experimental, psychological research; it is not the case that we can account for all of the results in this *Monograph*, or all of the results in the literature with our "cognitive-processing interpretation" of the effects of education.

Within the set of studies reported here, several discordant results need further study. We have already had occasion to note that there is no ready explanation for the absence of population differences in paradigmatic responding to nouns, while differences appear in response to verbs and adjectives. We have faulted "differential encoding" and "familiarity" explanations of these results with good cause. But we have no convincing explanation for the results. The most reasonable explanation draws on Moran's (1966) suggestion that paradigmatic responding represents subjects' attempts to find a response which bears a definable logical relation to the stimulus word (e.g., turkey-bird is an example of superordination, turkey-hen of subordination or subsumption in a single category). Such relations are perhaps more difficult to determine for adjectives (after one has responded "short" to "tall," what else is there to do?). If so, we might expect highly educated subjects to be more practiced in searching for such logical relations, the consequence of which is more paradigmatic responding. This interpretation, ad hoc as it is, has the virtue of connecting the free-association results with other features of the data, such as the educated subjects' reliance on the logical relations given within the problem as a basis for responding to syllogisms.

We were surprised by the lack of education-related differences in the replication of Birch and Bortner's attribute-matching study, especially when the matching set contained competing bases for a reasonable response. We strongly suspect that the failure to obtain differences by educational level resulted because, even when there were competing bases of solution, an answer that was scored as correct could be obtained by functional means, which the Mayan farmers were experts in producing.

The free-recall and short-term location-recall studies also produced an apparent discrepancy. In both tasks "primacy" is said to reflect the operation of rehearsal; only educated subjects demonstrated primacy in the location-recall task, but all groups demonstrated primacy in the (presumably more difficult) free-recall task.

Turning to the most closely related literature on education and cognitive development, we must consider the studies initiated by Kagan and his associates in neighboring Guatemala (Irwin et al. 1978; Kagan & Klein 1973; Kagen, Klein, Finley, & Rogoff 1976). Although concerned with broader cultural comparisons, the Guatemalan studies (which include free recall and recognition memory among other tasks) have been interpreted as evidence that cultural differences in basic cognitive processes are minimal by the time children have reached puberty, although substantial differences may exist in infancy.

Speaking to the free-recall results, we can see two major variables which differ between Kagan and Klein's study and our own: we presented materials for several trials (instead of a single presentation), and the range of educational experience of our subjects was far greater than in the study reported by Kagan and Klein. It would not be overlooked that we observed group differences only at the higher levels of education (primarily secondary school), and that with respect to category-based reorganization of the list during recall group differences appeared only after repeated trials with the list. For comparable education levels and amount of exposure to the experimental materials, our results are compatible with Kagen and Klein.

Speculative Uncertainties

Our "speculative" doubts arise not so much from unexplained experimental results as from very general doubts about the logic of the entire enterprise represented by our experiments and survey. Although our concerns with generalization from experimental data have been voiced previously (see Cole & Scribner 1974, 1975), they bear repeating in the face of seemingly massive confirmatory evidence that education fosters general cognitive development.

This conclusion rests upon the assumption that our experimental tasks are representative samples of cognitive problems people normally encounter and the processes which they apply to such problems. The problem of "representative design" is an old issue in psychology (see Brunswik 1955, pp. 656–750). It can be pushed into the background if our sole concern is with accounting for behavior in the situations we have chosen to observe, but it cannot be ignored when we want to assert that the behavior manifested in our experiments represents a general characteristic of the way people process information about the world.

It is perhaps easier to ask how strong our assumptions about test performance and cognitive development really are if we consider some examples from very different domains of behavior. Suppose, for example, that we wanted to assess the consequences of learning to be a carpenter. Sawing and hammering are instances of sensorimotor coordination; learning to measure, mitre corners, and build vertical walls requires mastery of a host of intellectual skills which must be coordinated with each other and with sensorimotor skills to produce a useful product (we are sensitive to this example owing to our own lack of success as carpenters!). To be sure, we would be willing to certify a master carpenter as someone who had mastered carpentering skills, but how strong would be our claim for the generality of this outcome? Would the measurement and motor skills learned by the carpenter make him a skilled electrician or a ballet dancer, let alone a person with "more highly developed" sensorimotor and measurement skills?

Lest it be thought that the example is too absurd to merit juxtaposition with the outcome of schooling, consider psychological experiments in light of the contexts from which their procedures have been derived and the domains in which they are routinely applied.

Some version of virtually every experimental task reported in this *Monograph* can be found in Alfred Binet's early work on the development of behavior samples which would predict children's success in school. The inspiration for their content came from an examination of the school curriculum, combined with Binet's sage guesses about the fundamental principles that underlay success in mastering that curriculum. The correlation between successful performance on Binet's tasks and success in school was a tautology; the items were picked because they discriminated between children at various levels of academic achievement. Might we not be witnessing the converse of that process when we observe people with educational experience excelling in experimental tasks whose form and content are like those they have learned to master in school? Is there any difference in principle between their excellence in recalling word lists and the master carpenter's ability to drive in nails quickly? After all, practice makes perfect; if we test people on problems for which they have lots of practice, why should we be surprised when they demonstrate their competence? Conversely, what leads us to conclude that they will be equivalently good at solving problems for which they have no specific practice?

The answer to this latter question, of course, is that we expect practice to be more than locally specific in its effects, because we expect it to transfer. In an earlier discussion of the cognitive consequences of schooling (Scribner & Cole 1973), we found that differences in transfer between problems was one of the distinguishing features of educated and noneducated adults. Unschooled populations tended to treat learning and memory problems as if they were discrete experiences, each constituting a new problem; there was a lack of learning to learn across problems which we are used to thinking of as examples of the same kind (even a lack of improvement from trial to trial within the same experiment). Schooled groups, by contrast, show marked learning across trials in free-recall studies, learning to learn, and a general tendency to treat classes of problems as instances of the same type, applying common operations in appropriate ways. This is just the kind of results we would expect if an outcome of formal education was the development of flexible problem-solving routines and rules for their application.

The weakness in this argument brings us right back to the problem of representative design. Perhaps the impression of educated subjects as general problem solvers is an illusion produced by the narrow range of tasks, all of them derived from school contexts, which we selected to represent the domain "cognitive development." The fact of the matter is that we have no direct evidence that educated subjects differ identifiably from their uneducated counterparts in the way they transfer their learning in any contexts other than our tests. Just as important, we have no idea of how often the intellectual demands represented by our experimental tasks are even encountered outside of the educational context from which they were derived.

When we step back from our close-up look at experimental results and take a more commonsense approach to our topic, it is easy to generate doubts on both scores.

Consider first the occasions upon which one is likely to meet demands such as those set up by our various tasks outside of an educational setting. How often is one asked to recall a set of 20 discrete items following a single exposure lasting only a second or 2 in duration for each item (the conditions obtained in our free-recall procedure)? Not only are such instances rare ("Hey, Charlie, name all the major league baseball teams," or "Who came to the club meeting last week?"), they are also treated in a more "negotiable" fashion. If the question is genuine (someone really wants to know who attended the meeting), the person required to recall would undoubtedly be able to jot down names if he could write, or ask his interlocutor for help ("Did I name Jones?"). We may try to rehearse an unfamiliar phone number, but it is more likely that we will jot it down, look it up in the phone book (if we can read and write), or call information. We often entertain the hypothetical ("What would have happened if I dropped a rock on your toe?"), but encountering a syllogism where the conclusion follows only according to logical rules contained within the problem seems a very rare event indeed; talk devoid of empirical content is something we are more likely to complain about than engage in.

There is virtually no evidence concerning the way educated and non-educated people respond to mundane intellectual demands of the sort we have just been imagining. One problem with gathering such evidence is that we would want to be certain that our observations were not contaminated by such factors as differential familiarity with the materials, amount of practice, and other factors which would influence our interpretation of the outcome. In other words, we would be tempted to set up an *experiment*. In so doing, we would, of course, undo the enterprise. A second problem arises from the fact that we know of no general procedures for specifying the similarity or difference between "mundane" and "experimentally contrived" tasks to warrant statements about performance differences between them. The purpose of an experiment is to insure that the analyst can know the demands of the task presented the subject; we have questioned both the adequacy of the specification obtained in this way and the inferential limitations which arise from the very act of creating a special, analyzable, task environment. But by the same token, when we do not artificially rig the task environment, we have little by way of tools for specifying what it might be. Ingenious attempts to overcome this difficulty are now under way using domains such as arithmetic to provide a means of nonexperiment-bound, yet specific intellectual tasks, but this line of research remains very much a problem for the future (see Ginsburg 1977; Lave 1977).

Some Consequences of a "Skills" Approach to the Effects of Schooling

The connection between identifiable work tasks and the consequences of education strikes us as a useful road to follow. When psychologists move from experimental data to assessment of policy questions, their claims for educational consequences (and the hoped-for changes by educational planners) are unlikely to depend upon proving that education produces intellectual development (although this belief crops up in a substantial portion of the literature linking literacy to economic development [see Gray 1956]). Rather, education is advocated as a means of instilling skills needed by a technological society, along with the values and aspirations needed to induce people to apply those skills. While we have to remain agnostic on the question of education's influence on helping someone to remember who attended a recent meeting, there are plausible grounds for believing that the information-processing skills which school attendance seems to foster could be useful in a variety of the tasks demanded by modern states, including clerical and management skills in bureaucratic enterprises, or the lower-level skills of record keeping in an agricultural cooperative or a well-baby clinic. These activities may well be facilitated by skills currently transmitted in schools.

If this is the case (as we believe it to be), the skills transmitted in schools do have applications in important nonschool settings. But until there is stronger evidence to the contrary, we advocate adoption of the rather restricted view of the consequences of education which we are adopting here. Like carpentry for a carpenter, it provides children with instruments that may be useful to them in specific work contexts. This perspective urges on us a reexamination of the *cognitive* basis for a relation between education and later income in terms of the respective skills required by the two settings.

Evidence about the determinants of people's current occupations gathered as part of our demographic survey emphasizes that in the Yucatan, like the United States, education is important for job selection. A series of regression equations showed two variables to be key determinants of informants' current occupational status (and, hence, income): highest grade, and sex. Other demographic factors (e.g., town modernity, family size, and languages spoken) had a lesser influence, though all in the expected directions. The impressive fact about the determinants of occupation is that with the exception of highest grade they all constitute accidents of birth from the informants' point of view. The importance of education as a source of social mobility among Yucatecans is a social fact.

But what is the basis of this fact? Are we observing a repetition of our experience in the United States, where educational requirements for jobs have increased to meet the current (increasing) level of educational attainment, irrespective of the relation between the skills transmitted in school and job-related skills? Or are we observing evidence that skills transmitted in schools do transfer to work domains? Until there is a substantial body of evidence demonstrating the generality of tested skills beyond school-like circumstances, and until their utility has been demonstrated through detailed studies of the similarity between the activities demanded in tests and various occupations, we must conclude that educational attainment cannot be legitimately used as a blanket prerequisite for employment in modern economic enterprises. Each job must be evaluated for the extent to which particular skills transmitted in school are necessary prerequisites.

One implication of this work for psychological theory seems clear, although it will be important primarily to those who are actively engaged in the research: developmental, cognitive research in the United States and other industrialized countries, where years of education and age go hand in glove, has been studying the consequences of education rather than culture-free developmental laws. A great deal remains to be done to disentangle the various knots which currently bind this line of inquiry.

APPENDIX

QUESTIONS ON EDUCATIONAL DEMOGRAPHIC SURVEY

1. Have you studied (been to school)? Why did you stop studying (or fail to go to school)?
2. Can you read and write?
3. Have you repeated a grade? If yes, how many times?
4. Do you have any brothers and sisters? If yes, what is their sex, age, years of education?
5. Did your father go to school? How old is he and how many years of schooling does he have?
6. Did your mother go to school? How old is she and how many years of schooling does she have?
7. Do you have children? What are their ages and how many years of schooling do they have?
8. (If you have children) do you think that it is important that your children study?
   a) It is not important
   b) It is important for some children
   c) It is important for all
   d) It is very important for all
9. Do you think that it is necessary to go to school to do well in Mexico? Why?
10. Do (did) your parents think that it is (was) important to go to school?
    a) Not important
    b) Perhaps it (was) important
    c) It is (was) important
    d) It is (was) very important
11. What do your brothers and sisters think?

12. What (minimum) grade of study do you think is necessary for a man to live well in Mexico?
    a) Some primary
    b) Primary
    c) Secondary
    d) Technological, preparatory, normal
    e) Professional
    f) More than professional
13. Do you think that a woman ought to have the same studies as a man?
14. What languages do you speak? What was your first language?
15. What language do you speak at home?
16. Do you think that a woman has the same opportunity to study as a man in Mexico? Why?
17. Was there a school in your town? WHAT GRADE DID IT GO TO? How far was it to (from) your house? Who paid your school expenses?
18. What work does (did) your father do?
19. What work do you do?
20. Have you had other jobs? What were they?
21. Did you help your parents when you were studying? Did you miss classes? If yes, how many times?
22. Have you traveled to or lived in other towns? Which were they?
23. How much money do you think it is necessary to earn daily to live well in Mexico?
24. What work would you like to do to earn this money?
25. Have you helped other members of your family to go to school? How many and to what grade?
26. Where would you like to live if it were possible?
27. Would you like to continue studying (or go to school)?
28. What profession would you like to have?
29. What do you want to do when you finish studying?
30. What is the most important reason for which students do not continue studying?
    a) Problems with their families
    b) Because there is no school near
    c) Responsibilities that make it impossible to go
    d) Money for their expenses
    e) The opinion of their family or friends
31. Some think that only the rich can study. What do you think?
32. What do you think of those that do not continue studying after they finish their primary education?
33. Some say that school is responsible for problems in the family. What do you think?

34. Do you believe that those that have primary education have more opportunity than those that do not?
35. What do you think is the most important for the future of your country?
    a) The hard work of the people
    b) The good plans of the government
    c) God's help
    d) Good luck
36. Have there been times in the past when you have thought a lot about a national problem such as social security for the aged and children and want to do something?
    a) Never
    b) Occasionally
    c) Often
37. If you were to meet a person that lived in another country many kilometers from here, do you think that it would be possible to understand the way in which he thinks?
38. Which are the problems that you think that the country has today?
39. How often do you read the news in the paper?
    a) Every day
    b) Every week
    c) From time to time
    d) Never
40. Professors and teachers in universities are studying things like the factors that determine if a child is a boy or a girl and the way in which a seed changes into a plant. Do you think that these investigations are:
    a) Very good
    b) At times good
    c) At times bad
41. Two boys of twelve years of age were working on a farm, and while they worked they were talking about how to grow more corn in less time. The two decided to ask their fathers about the problem when they returned home. That night the father of one boy said that their thoughts on the matter were a good thing and that it was very important to give consideration to problems such as growing more corn in less time. The father of the other said that the way to grow corn was the way in which it had always been done and that things like changing it would be only to lose time and nothing more.
    Which father do you think said the wisest words, the first or the second?
42. WAIS: I have here a group of words that are in pairs and I want you to tell me in which way they appear alike or if they do not appear alike or have no relation [presented in this order].
    1) pineapple-banana
    2) corn-squash

3) machete-hoe
4) pot–frying pan
5) horse-goat
6) watermelon-potatoes
7) woman-boy
8) man-horse
9) honey-water
10) fly-tree
11) gasoline-wood
12) bicycle-automobile
13) cedar-mahogany
14) air-water

REFERENCES

Berlin, B.; Breedlove, D.; & Raven, P. General principles of classification and nomenclature in folk biology. *American Anthropologist*, 1973, **75**, 214–241.

Birch, H. G., & Bortner, M. Stimulus competition and category usage in normal children. *Journal of Genetic Psychology*, 1966, **109**, 195–204.

Brown, A. L. The role of strategic behavior in retardate memory. In N. R. Ellis (Ed.), *International review of research in mental retardation*. Vol. **7**. New York: Academic Press, 1975. (a)

Brown, A. L. The development of memory: knowing, knowing you don't know and know how to know. In H. Reese (Ed.), *Advances in child development and behavior*. Vol. **10**. New York: Academic Press, 1975. (b)

Brown, A. Development, schooling and the acquisition of knowledge about knowledge. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, N. J.: Erlbaum, 1977.

Brunswik, E. The conceptual framework of psychology. *International encyclopedia of unified science*. Vol. **1**, pt. 2. Chicago: University of Chicago Press, 1955.

Cole, M. An ethnographic psychology of cognition. In R. W. Brislin, S. Bochner, & W. J. Lonner (Eds.), *Cross-cultural perspectives on learning*. Beverly Hills, Calif.: Sage, 1975.

Cole, M.; Gay, J.; Glick, J. A.; & Sharp, D. W. *The cultural context of learning and thinking*. New York: Basic, 1971.

Cole, M., & Scribner, S. *Culture and thought*. New York: Wiley, 1974.

Cole, M., & Scribner, S. Theorizing about socialization of cognition. *Ethos*, 1975, **3**, 250–268.

Dasen, P. Introduction in *Piagetian psychology*. In P. R. Dasen (Ed.), *Cross-cultural contributions*. New York: Gardner Press, 1977.

Entwistle, D. R. Developmental sociolinguistics: a comparative study in four subcultural settings. *Sociometry*, 1966, **29**, 67–84.

Fjellman, J. The myth of primitive mentality: a study of semantic acquisition and modes of thought. Unpublished Ph.D. dissertation, Stanford University, 1971.

Flavell, J. H., & Wellman, H. M. Metamemory. In R. V. Kail, Jr., & J. W. Hagen, *Perspectives on the development of memory and cognition*. Hillsdale, N.J.: Erlbaum, 1977.

Gay, J., & Cole, M. *The new mathematics and an old culture*. New York: Holt, Rinehart & Winston, 1967.

Ginsburg, H. Some problems in the study of schooling and cognition *Quarterly Newsletter of the Institute for Comparative Human Development*, 1977, **1**, 7–10.

Gray, W. *The teaching of reading and writing*, *UNESCO Monographs on Fundamental Education* (Paris). 1956 (No. 10).

Greenfield, P. On culture and equivalence. In J. S. Bruner, R. R. Olver, P. M. Greenfield, et al., *Studies in cognitive growth.* New York: J. Wiley, 1966.

Greenfield, P. Comparing dimensional categorization in natural and artificial contexts: a developmental study among the Zinacantecos of Mexico. *Journal of Social Psychology*, 1974, **93**, 157–171.

Greenfield, P., & Bruner, J. S. Culture and cognitive growth. *International Journal of Psychology*, 1966, **1**, 89–107.

Greer, C. *The great school legend.* New York: Basic, 1972.

Hagen, J. W., & Hale, G. A. The development of attention in children. In A. D. Pick (Ed.), *Minnesota symposium on child psychology.* Vol. **7**. Minneapolis: University of Minnesota Press, 1973.

Hall, J. Verbal behavior as a function of amount of schooling. *American Journal of Psychology*, 1972, **85**, 277–289.

Harmon, D. *Community fundamental education.* Lexington, Mass.: Lexington Books, 1974.

Holtzman, W. H.; Diaz-Guerrero, R.; & Swartz, J. D. *Personality development in two cultures.* Austin: University of Texas Press, 1975.

Hudson, R., & Dunn, J. A major modification of the Bousfield (1966) measure of category clustering. *Behavioral Research Methods & Instrumentation*, 1969, **1**, 110–111.

Hunt, J. Mc.V. *Intelligence and experience.* New York: Ronald, 1961.

Irwin, M.; Engle, P. L.; Yarbrough, C.; Klein, R. E.; & Townsend, J. The relationship of prior ability and family characteristics to school attendance and school achievement in rural Guatemala. *Child Development*, 1978, **49**, 415–427.

Irwin, M. H.; & Mclaughlin, D. H. Ability and preference in category sorting by Mano schoolchildren and adults. *Journal of Social Psychology*, 1970, **82**, 15–24.

Jencks, W. *Inequality.* New York: Basic, 1972.

Jensen, A. How much can we boost I.Q. and scholastic achievement? *Harvard Educational Review*, 1969, **39**, 1–123.

Jensen, A. R. The role of verbal mediation in mental development. *Journal of Genetic Psychology*, 1971, **118**, 39–70.

Johnson, S. C. Hierarchical clustering schemes. *Psychometrika*, 1967, **32**, 241–254.

Kagan, J., & Klein, R. E. Cross-cultural perspectives on early development. *American Psychologist*, 1973, **28**, 947–961.

Kagan, J.; Klein, R. E.; Finley, G. E.; & Rogoff, B. A study in cognitive development. *Annals of the New York Academy of Science*, 1976, **285**, 374–388.

Kent, G. H., & Rosanoff, A. J. A study of association in insanity. *American Journal of Insanity*, 1910, **67**(1 and 2), 37–96.

Lave, J. Tailor-made experiments and evaluating the intellectual consequences of apprenticeship training. *Quarterly Newsletter of Institute for Comparative Human Development*, 1977, **1**, 1–3.

Luria, A. R. Rech' i intellekt. Derevenskogo, gorodskogo i bespirizornogo rebenka. Gosudarstvennoe izdatel'stvo RSFSR, 1930. Reprinted in *Soviet Psychology*, 1974, **13**, 5–6.

Luria, A. R. *The historical development of cognitive processes.* Cambridge, Mass: Harvard University Press, 1976. (Originally printed in Moscow, 1974.)

McClaran, M. Some lexical and syntactic structures in Yucatec Maya. Unpublished Ph.D. dissertation, Harvard University, 1971.

Maccoby, E. E., & Hagen, J. W. Effects of distraction upon central versus incidental recall: developmental trends. *Journal of Experimental Child Psychology*, 1965, **2**, 280–289.

Moran, L. J. Generality of word-association response sets. *Psychological Monographs*, 1966, **80**(4, Whole No. 612).

Murdock, B. B., Jr. *Human memory.* Hillsdale, N.J.: Erlbaum, 1974.

Pillsbury, W. B. Selection—an unnoticed function of education. *Scientific Monthly*, 1920, **12**, 62–74.

Rohwer, W. D., Jr. Elaboration and learning in childhood and adolescence. In H. W. Reese (Ed.), *Advances in child development and behavior*. Vol. **8**. New York: Academic Press, 1973.

Rosenzweig, M. R. Word associations of French workmen: comparisons with associations of French students and American workmen and students. *Journal of Verbal Learning and Verbal Behavior*, 1964, **3**, 57–69.

Scribner, S. Situating the experiment in cross-cultural research. In K. F. Riegel & J. A. Meacham (Eds.), *The developing individual in a changing world*. Vol. **1**. Chicago: Aldine, 1976. (a)

Scribner, S. Modes of thinking and ways of speaking: culture and logic reconsidered. In R. O. Freedle (Ed.), *Discourse production and comprehension*, Hillsdale, N.J.: Erlbaum, 1976. (b)

Scribner, S., & Cole, M. Cognitive consequences of formal and informal education. *Science*, 1973, **182**, 553–559.

Sharp, D. W., & Cole, M. Patterns of responding in the word associations of West African children. *Child Development*, 1972, **43**, 55–65.

Stefflre, V.; Reich, P.; & McClaran, M. Some eliciting and computational procedures for descriptive semantics. In P. Kay (Ed.), *Explorations in mathematical anthropology*. Cambridge, Mass.: MIT Press, 1971.

Stefflre, V.; Vales, V.; & Morley, L. Language and cognition in Yucatan: a cross-cultural replication. *Journal of Personality and Social Psychology*, 1966, **4**, 112–115.

Stolz, W., & Tiffany, J. The production of "child-like" word associations by adults to unfamiliar adjectives. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 38–46.

Thurstone, L. I. *Primary Mental Abilities*. Chicago: University of Chicago Press, 1938.

Wagner, D. A. The development of short-term and incidental memory: a cross-cultural study. *Child Development*, 1974, **45**, 389–396.

Winer, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

Young, F., & Torgerson, W. Torsca, a Fortran IV program for Shephard-Kruskal multidimensional scaling analysis. *Behavioral Sciences*, 1967, **12**, 498.

COMMENTARY AND REPLY

COMMENTARY BY HERBERT P. GINSBURG

This *Monograph* stems from a relatively new area of research which investigates cognitive development in cross-cultural context. The *Monograph* addresses two main issues, one substantive and the other methodological. First, how do schooling, age, and culture affect the development of cognitive functioning? Second, what factors determine attendance at school? To the second question, the writers give as clear an answer as is possible under the messy conditions of cross-cultural research. Most important, they show that IQ plays only a minor role in the selection of children for schooling. But, as the writers themselves acknowledge, the research fails to provide conclusive answers to the first question. Nevertheless, the *Monograph* opens up new directions for theory and research. The main aim of this comment is to examine these possible new directions.

The Research

With respect to the main substantive question, the results are inconsistent and difficult to interpret. (As mentioned above, I do not refer to the demographic survey which is extremely useful and which I do not discuss further.) The empirical findings are complex: some tasks reveal differences correlated with schooling; other tasks do not. For example, in the word-pair similarity tests (see table 9), the group differences on the proportion of correct responses are relatively small. Yet, an alternative method of scoring the data (the average WAIS score) does yield reliable differences. Given data like these, the interpretation must also be complex. Thus: " . . . all populations studied *can* make verbal classifications on the basis of shared semantic features. But as the task takes on extra properties . . . semantic attributes may not be used, especially by the less-educated subjects . . . " Also, the results may hinge on minor methodological variations: " . . . the estimated magnitude of the effect depends upon details of procedure . . . " (p. 34).

To explain these results, the writers propose that differences between educated and noneducated subjects depend mainly on the extent to which the task is structured. "If the task itself organizes subjects' responses so that it can be solved without recourse to special intellectual work, or if the task's demands are beyond the reach of the educated subjects, little or no education-related differences in performance are to be expected" (p. 77). In other words, if a task "tells" the subject what to do, the poorly educated person generally has the competence to do what is required. But if the task is relatively open-ended, the uneducated person may not be able to determine which of his available competencies to use. This post-hoc hypothesis is a reasonable but not wholly satisfactory account of the writers' data. It accounts for many, but not all, of the results. As the writers themselves point out, they have no way of explaining the absence of group differences in the paradigmatic responding to nouns. Furthermore, it seems self-evident that the hypothesis cannot be true for intellectual performance in general. For example, it must be the case that uneducated subjects familiar with, say, navigation or car repair or hunting would require less task structure to perform the appropriate cognitive operations than would educated subjects unfamiliar with these areas. Finally, we would add that it would have been useful if the authors had discussed their data in light of previous hypotheses concerning the role of schooling. For example, Scribner and Cole (1973) proposed that uneducated people tend to solve individual problems singly, whereas educated people tend to solve them by a general rule. They also propose that educated subjects are more able to use language to describe the tasks and their own activities than are uneducated subjects. How do the present data shed light on these hypotheses?

Ecological Validity

The complexity of the data, the failure of the proposed hypothesis to account convincingly for all of them, and various speculative uncertainties all combine to convince the writers that the entire enterprise needs to be reevaluated. Even if the results were to be replicated, and even if the hypothesis has some degree of validity, what is the basic meaning of the research? "Are we dealing with a circumscribed set of tasks which have little significance outside of experiments (and possibly schools)?" (p. 79). According to the writers, the key issue is the "ecological validity" of tasks. Essentially, the question is whether the experimental tasks employed in this study, and indeed in the great majority of cross-cultural researches, reveal anything fundamental concerning cognitive activity in everyday situations. "Perhaps the impression of educated subjects as general problem solvers is an illusion produced by the narrow range of tasks, all of them derived from school contexts. . . . The fact of the matter is that we have no direct evidence that educated subjects differ identifiably from their uneducated counterparts in the way they transfer their learning in any contexts other than our tests. Just as important, we have no idea of how often the intellectual demands represented by our experimental tasks are even encountered outside of the educational context from which they were derived" (p. 82). The question of "ecological validity" is indeed fundamental, and not only for cross-cultural research: similar doubts may arise in connection with much domestic research employing many of the same test procedures whose significance is at issue in cross-cultural study.

The issue of "ecological validity" warrants careful analysis. Consider first the writers' claim that researchers can say little concerning everyday cognition since they use a " . . . narrow range of tasks, all of them derived from school contexts." Our knowledge of everyday cognition is indeed limited, but school-derived tasks are only one reason for our ignorance. More important, the tasks typically employed in cognitive developmental research are not appropriate for the investigation of individual differences within stages, cultural variation, and related noncognitive factors like motivation.

As the writers point out, school-based tasks—like the IQ test, or the similarity scale used in this *Monograph*—were originally developed by Binet and others to predict academic achievement in Western schools. Such tests are indeed useful for that very purpose: with reasonable accuracy, they do predict later performance on similar tests which are given in schools. While such tests have reasonable predictive validity, they tell us very little about cognitive functioning in general. What little we know about "intelligence" does not derive from the study of IQ tests. Further, the tests are typically "culturally unfair" in the sense that they are not designed to reveal the unique cognitive strengths of culturally distinct groups. The tests do accurately predict how such people will do in school—usually badly, which is no great revelation—but the tests say little of value about intellectual functioning, and hence are inappropriate for use in cross-cultural research concerned with everyday cognition, or indeed in any research concerned with cognitive process.

But cross-cultural research often uses other tasks whose "ecological validity" is in doubt. These are tasks which derive from experimental child psychology and from Piaget. In the former category are tasks like transposition, reversal shifts, and paired associates. In the latter are conservation, seriation, classification, etc. In general, these tasks carefully control extraneous factors so as to provide insight into fundamental cognitive processes. The aim is to study the "general human mind" (to use Wundt's phrase) at its several stages of development. Thus, the tasks are used to determine whether the processes of memory change with development or to understand the structure of the concrete operations.

Are such tasks "ecologically valid," and can they play a useful role in cross-cultural research? In one sense, many of the tasks are ecologically valid in that they measure basic cognitive process, or aspects of the general human mind, which in some ways must be employed in everyday life. Thus, Piaget's seriation task must be related to everyday attempts to deal with quantitative relations in the real world. Indeed, some of the tasks may have originated in the observation of cognitive activities in everyday life. For example, the transposition task may have derived from the observation of Gestalt psychologists and their predecessors that melodies are easily and obviously transposed across scales. Similarly, Ebbinghaus developed his memory tasks in an attempt to simplify and gain control over what he considered to be everyday processes of memory. So the experimental and Piagetian tasks are "ecologically valid" in the sense that they refer to basic cognitive processes which must be used in the real world.

Tasks of this nature, perhaps slightly modified to accommodate local cultural variations, may be used profitably in cross-cultural research to shed light on the issue of "cognitive universals"—whether the "general human mind" is indeed general across cultural variation. Thus, experimental tasks can provide useful information concerning whether, for example, the emergence of transposition is a stable developmental phenomenon across cultures or whether members of all societies achieve the stage of formal operations.

At the same time, experimental and Piagetian tasks suffer from severe limitations with respect to the cross-cultural study of everyday cognition. For one thing, the tasks are not designed to elucidate individual differences within stages. In the case of transposition, the experimentalist is interested in the differences between those who transpose and those who do not, not in the differences among transposers or nontransposers. Or the Piagetian wants to know whether subjects have achieved the stage of formal operations or are still limited to concrete operations, but is not interested in how formal operational adolescents differ from one another. Yet in the study of everyday cognition within the context of culture one is directly interested in individual differences within stage: one wants to know, for example, how formal operations are used by members of one profession or linguistic group as opposed to another. Or one wants to know how adults in various cultures differ with respect to their memory processes as used in everyday situations. The experimental and Piagetian tasks cannot be used in their present form to conduct such investigations.

Second, since the experimental and Piagetian tasks are designed to eliminate "extraneous factors" in order to focus on *competence* in basic process, they cannot be sensitive to variables of culture, language, and motivation which are at the heart of everyday cognition. What is "extraneous" in one case is central in the other. Thus, to investigate competence in transposition one wishes to eliminate difficulties in understanding language, or differential familiarity with materials, or lack of motivation. But these are precisely the variables of interest in cross-cultural research in everyday cognition: how do differences in language, or familiarity with cultural objects, or motivational variations related to culture, affect cognitive *performance*? Indeed, the experimental and Piagetian tasks almost demand a special type of subject—namely, middle-class European or American children who can be counted on to display proper motivation, to interpret instructions properly, and in general to act as cooperative subjects. This is entirely reasonable: subjects other than these may only produce "noise" insofar as the basic aim of the research is concerned. Such research depends upon a special class of subjects in much the same way as early introspective research depended on trained observers. Yet these are not necessarily the subjects of particular interest to cross-cultural research.

Third, although the basic processes measured by the experimental and Piagetian tasks must bear some relation to everyday cognition, the links between the two areas are now obscure. In general, we have failed to conduct serious investigations of the relations between the tasks and their possible analogues in everyday life. To a very large extent, the laboratory tasks have taken on a life of their own, and the investigator tends to forget their origins and to ignore their possible everyday significance or lack of it. Thus, while the seriation task may reflect something basic concerning quantitative thought, we do not know from Piaget's theory how such thought is used in everyday life, nor can we use the seriation task to study the issue.

In brief, school-related tasks are inappropriate for our research purposes since they have so little to say about cognition. Experimental and Piagetian tasks are useful for the investigation of cognitive universals, but suffer from severe limitations with respect to the study of individual differences within stages, and the influence of motivation, language, and culture on cognitive performance. Further, these tasks have taken on a life of their own; the theories associated with them provide little insight into everyday cognition, and the tasks themselves cannot easily be used to study it.

New Directions

It seems clear then that the current tools of cognitive research are generally inappropriate or inadequate for the study of cognition in everyday life. We obviously require new approaches; we have to go beyond the usual experiments. One solution, as the writers of the present *Monograph* suggest, is to engage in naturalistic research designed to shed light on the cognitive processes employed in everyday life. This gets us away from the study of the general human mind and Western school-related tasks and turns us toward what Wundt called "ethnic psychology." Such a psychology is not concerned only with competence—that is with what people can do under optimal circumstances. Instead it is concerned with what people actually do, under various conditions of culture, education, motivation, and personality. Thus, to understand, say, mathematical thinking in everyday life, one needs to know not only about the concrete operations (the child's competence) but about other cognitive operations that are used, why he uses or needs to use mathematical thinking at all, what role it plays in his culture, what it gets him, and what forms it takes within the local culture.

At present there are few examples of such research: its future success is not assured. Because of our long-term isolation in the laboratory and the limits of the Piagetian clinical interview, we have developed few theoretical models or research techniques appropriate for the study of everyday cognition. Yet the situation is beginning to change. Important attempts are being made, from different perspectives, to define an ecologically based psychology (Bronfenbrenner 1979; Neisser 1976a); and innovative research in the area is being conducted (e.g., Lave 1977). At present we do not know how successful these efforts will be or exactly what form an ecologically based cognitive psychology will take. It does seem clear, however, that such a psychology will look very different from current theoretical accounts. A cognitive psychology of everyday life—with or without a cross-cultural emphasis—must consider much more than the structures of logical thought or the mechanics of information processing. The new cognitive theory will be more *psychological* than the old: it must include a theory of performance, of culture, and of personality. The study of ordinary cognition is too important to be left to cognitive psychology alone.

Schooling

Just as we need to rethink our theory and measures of cognition, so we need to reconsider the "independent variable" of schooling. Typically, researchers—including the writers of the present *Monograph*—have treated schooling as a quantitative variable, using numbers of years in school as the chief measure. But this is only a crude approximation: as researchers surely recognize, schooling is an extremely complex experience. Research needs to go beyond mere consideration of years of schooling as a variable. Schooling, which accounts for a major proportion of the child's life, involves many different kinds of learning and cognitive activities. It also involves, to a significant extent, training in values, moral education, social learning, self-exploration, and a variety of activities not usually considered under the rubric of "cognition." Psychologists know very little about what actually happens in schools. To be sure, there is some recent ethnographic work in this area (see, e.g., Ogbu 1978), but psychologists are ordinarily unfamiliar with the day-to-day workings of schools, with the textbooks children use, and with what they learn. To examine "the effects of schooling," psychologists need to do "field work" in the schools and need to develop reasonable theoretical accounts of the school ecology and of the activities taking place there.

It is especially important to understand the structure of academic knowledge. One classical argument asserts that a major benefit of education is that school learning transfers to new situations. Yet until we know something of what children actually learn in school, it is hard to study such transfer. Fortunately, the topic of academic knowledge has recently attracted the interest of cognitive psychologists, who are now beginning to make progress in this area (see Anderson, Spiro, & Montague 1977).

Usually, like the writers of the present *Monograph*, researchers assume that schooling has generally beneficial effects on cognitive development. But it is important to consider the possibility that schooling may have no effects or even negative effects on intellectual development. In regard to the former, there is some evidence that schooling may result in knowledge which is severely limited to specific contexts, which bears little relation to the child's ordinary experience, and which cannot easily be transferred (Ginsburg 1977). Similarly, Wertheimer (1945), in a famous observation, points out that students' solutions to geometry problems often lack insight, are "blind," and cannot be generalized even within the relatively narrow context of similar geometry problems. Indeed, since schooling is often seen to result in narrow and limited forms of learning, critics of the educational system sometimes propose that to foster education children should be removed from the schools and placed in apprenticeship situations where they might be more able to appreciate intellectual activity's relevance for reality. (In this regard, see Neisser's [1976b, pp. 135–146] discussion of "academic intelligence," a rather limited form of intellectual activity which happens to be favored by academics.)

Not only may schooling be irrelevant, limited, and trivial, it may also produce distinctly negative effects. One common outcome of schooling is to produce in the schooled a fear of what they have studied. Thus, mainly as a result of schooling, math anxiety is extremely widespread; many students, long after they have left school, retain a distinct bias against literature and poetry; a good many students after instruction will have nothing whatsoever to do with the physical sciences. In short, many schools produce an extreme dislike for anything intellectual. To a large extent, successful work in literature, mathematics, or science requires the systematic undoing of what is taught in school.

We see then that our ignorance concerning schooling and what is learned there is great, and that we should be alert to the possibility that in some ways schooling may have negligible or even detrimental effects on intellectual development.

As the writers point out, research should examine the relations between schooling and work activities. It has often been assumed that schooling provides individuals with intellectual skills necessary for success in various occupations. Yet, whether this is true and, if so, how the process works in detail is by no means clear. In this regard, it might be helpful to examine noncognitive skills which may be fostered in schools, such as attentiveness, obedience, and persistence.

Finally, we need to consider education in its most general sense. We need to consider how schooling may involve more than specific learning or applications of particular skills. We need to understand schooling as the inculcation of culture, style, and values. Education may be conceptualized as a "general context for knowledge" (Broudy 1977, pp. 1–18). When successful, education does more than provide skills: it stimulates the imagination, provides respect for knowledge and an appreciation of beauty, gives insight into the excitement of intellectual work. Psychologists need to widen their perspective to include these central elements of a humane education.

References

Anderson, R. C.; Spiro, R. J.; & Montague, W. E. (Eds.). *Schooling and the acquisition of knowledge*. Hillsdale, N.J.: Erlbaum, 1977.

Bronfenbrenner, U. *Experimental human ecology*. Cambridge, Mass.: Harvard University Press, 1979.

Broudy, H. S. Types of knowledge and purposes of education. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, N.J.: Erlbaum, 1977.

Ginsburg, H. *Children's arithmetic: the learning process*. New York: Van Nostrand, 1977.

Lave, J. Tailor-made experiments and evaluating the intellectual consequences of apprenticeship training. *Quarterly Newsletter of the Institute for Comparative Human Development*, 1977, **1**, 1–3.

Neisser, U. *Cognition and reality*. San Francisco: Freeman, 1976. (a)

Neisser, U. General, academic, and artificial intelligence. In L. B. Resnick (Ed.), *The nature of intelligence*. Hillsdale, N.J.: Erlbaum, 1976. (b)

Ogbu, J. U. *Minority education and caste*. New York: Academic Press, 1978.

Scribner, S., & Cole, M. Cognitive consequences of formal and informal education. *Science*, 1973, **182**, 553–559.

Wertheimer, M. *Productive thinking*. New York: Harper, 1945.

[**Herbert P. Ginsburg** (Ph.D. 1965, University of North Carolina) is a professor of psychology at the University of Maryland, Baltimore County, and director of the Program in Applied Developmental Psychology. Professor Ginsburg has been working on cross-cultural differences in mathematical thinking of children from the Ivory Coast of Africa, and is currently studying social class differences in mathematical thinking of children from the United States. Some of his current publications include *Childrens' Arithmetic* (New York: Van Nostrand, 1977), and *Piaget's Theory of Intellectual Development* (2d ed.; Englewood Cliffs, N.J.: Prentice-Hall, 1979) which was coauthored with Sylvia Opper.]

* * *

COMMENTARY BY ANN L. BROWN AND LUCIA A. FRENCH

For many years, Cole and his colleagues have carried out research investigating the ability of people in traditional cultures to perform cognitive tasks which Western educators and psychologists consider to be measures of cognitive status. The guiding principle governing these endeavors has been to determine how much of what we consider normal cognitive development is in fact an age-dependent developmental phenomenon, or to what extent it reflects the result of experiences associated with the degree and extent of formal schooling. To address such questions it is necessary to study both schooled and unschooled people who share similar backgrounds; to do so one must turn to cultures where schooling exists but is not universal.

The research reported in this *Monograph* gives a clear-cut answer to the original question: schooled population overwhelmingly outperform the unschooled on a variety of laboratory tasks traditionally used to assess intelligence or cognitive maturity. However, the investigators themselves seem less than happy with the interpretation that can and will be placed on their findings. Thus, we have in this *Monograph* the rather unusual situation where the conclusion does not follow simply from the data base; rather, it is laid over the data almost as a warning not to be taken too seriously. Many of our reactions are not to the body of the paper itself but to this conclusion. While we suspect that the conclusions drawn may be correct, we will adopt the role of devil's advocate and argue that the necessary information to support the conclusion is not yet in. We will begin, however, by considering the real contribution of the data provided by the Yucatan project before proceeding to a consideration of the wider implications drawn in the concluding section of this manuscript and expanded upon in other recent publications of Cole and his colleagues (Cole 1978; Cole, Hood, & McDermott 1978b; Cole & Laboratory for Comparative Human Cognition 1978a, 1978b).

The Effects of Schooling

There can be no doubt that the Yucatan project has provided invaluable information for those interested in mapping the course of cognitive development. The project is one of the most extensive and carefully controlled cross-cultural studies, rivaled only by the Liberian projects which were also pioneered by this group. The careful attention paid to socioeconomic factors such as family size, first language, sex, parental education, etc., is unique. Because of this care, and the fortuitous distribution of educational opportunities in the Yucatan, it is possible for Sharp, Cole, and Lave to conclude that the difference found between educated and uneducated Yucatecs were brought about by exposure to formal education rather than being the outcome of gross selection factors that determined the degree of schooling in the first place. To our knowledge, this is the first such study where it is possible to confidently rule out massive selection artifacts in the comparison of schooled and unschooled samples.

Another substantial contribution of the Yucatan project is the cohesive pattern of cognitive skills that emerges as the outcome of formal schooling. In a wide variety of traditional laboratory tasks, the educated Yucatecs differed from their uneducated counterparts in clearly diagnostic ways; that is, the educated did not always outperform the uneducated, but when differences did emerge it was in situations where some active transformation of the materials was necessary for effective performance. As Sharp et al. put it, "It's what you do with what you have that counts"; for the more highly educated participants had a greater predisposition to "engage in intellectual activities which are not rigidly predetermined by the structure of the task and which promote efficient performance. It is not differences in the information about the stimuli per se but differences in what people do with commonly available information that is critical." Education influences the readiness of people to use the information they have to help them to remember or to solve problems more efficiently. This pattern of results can readily be assimilated into current models of cognitive development and clearly supports the growing suspicion that much of what we regard as normal cognitive development—that is, the acquisition of a repertoire of effective, flexible strategies for coping with decontextual, closed-system problems—is very much affected by significant exposure to formal schooling.

Specific and General Skills

Sharp et al. seem to have serious reservations about whether the data they have so painstakingly collected and analyzed have any relevance to (*a*) important forms of everyday cognition (Cole et al. 1978b), and (*b*) policy making in traditional societies (Cole 1978). To address these questions we would like to emphasize a distinction between specific and general skills, a distinction made somewhat opaquely by the Cole group in recent publications. One conclusion often drawn is that schooling, like tailoring or carpentry, promotes certain *task-specific* skills. The particular skills engendered by school experience involve the ability to apply a specific set of strategies to the task of solving decontextualized academic problems commonly encountered in school curricula. But the tasks used by cognitive psychologists to measure the effects of formal schooling are variants of the very tasks selected by Binet and others precisely because they predict academic (school) success; their success at predicting is because they are representative of the type of tasks regularly encountered in school. As a result, the tasks on which the educated excel are those in which they, and they alone, receive practice at school. As Sharp et al. point out, the subsequent finding of a schooling advantage is tautological given the tasks selected to measure the putative benefits. The authors then are worried that the impression to emerge from projects like the Yucatan studies, of educated people as general problem solvers, may be an illusion. We lack the necessary information concerning whether the schooled populations will outperform the uneducated in situations not representative of school experience but more in keeping with "everyday" reasoning demands. In short, is it the case that schooling promotes specific skills relevant to schools but to nothing else?

At another level, Cole and his associates have argued that schooling influences modes of thinking in a much more general way, for by "giving children rules they can apply to many kinds of information, schooling may promote a special kind of 'learning to learn' so that procedures for processing new information become second nature" (Cole 1978, p. 53). If schooling really does enhance the ability to acquire new information and solve novel problems, in general, and not just in the school setting, it would surely be regarded as profoundly influencing its recipients. But we are caught in a bind, clearly recognized by Cole and his associates, for the very data base used to support the position that schools teach people how to learn and how to transfer solutions across task boundaries is information gleaned exclusively from performance on school-relevant tasks that tap schooled skills.

It is of major interest to establish whether schooling leads to a generalized ability to solve new problems, regardless of contexts, but there are practically no data to illuminate this point. A pessimistic note has been struck by the oft-quoted study of tailoring skills conducted by Lave (1977). She examined transfer of arithmetic and measurement skills by Liberian tailors with varying amounts of tailoring experience and varying degrees of formal education. Transfer problems were formally equivalent in terms of the underlying computational requirements but were presented in a surface format that was more appropriate either to the context of tailoring or to the context of formal education. Tailoring experience led to success across tailoring problems, whereas experience with formal education led to transfer across traditional school problems. Transfer operated within limited contexts and depended on specific experiences. These data do tend to support the specific-skills position, that formal education supplies school-relevant skills just as tailoring experience induces tailoring-specific skills and that any learning to learn advantage is context bound. Excellent as the Lave study may be, it is a very limited data base upon which to bolster the far-reaching conclusions concerning the limited affects of education hinted at in the conclusion of the *Monograph.*

Metacognition and Expertise

We would like to argue that despite the enormous effort and care invested in the Yucatan project, the cross-cultural data base is not only limited but dated. The Yucatan project was initiated in the early seventies and is fairly representative of the state of the art in cognitive development research at that time. The tasks selected for study are typical of the then popular laboratory enterprises, and the level of analysis is focused on the use of specific strategies to solve such problems—rehearsal, categorization, and the use of knowledge of general classes to solve problems. There is no direct consideration of the general factors involved in the control and orchestration of problem-solving strategies, general factors that have come to be subsumed under the heading "metacognition" (Flavell & Wellman 1977). Metacognition refers to the awareness about and conscious control of one's own cognitive processes, to the mental processes which can be broadly termed checking and monitoring activities. In a sense, what is involved is a general "keeping an eye on" one's cognitive operations to determine when strategic attempts to learn are called for, when failures to understand have occurred, whether it is likely that a problem can be solved on the basis of old knowledge, and, if not, what new knowledge must be sought, etc.

It seems likely that such general problem-solving knowledge will help in many situations outside of school where one is called upon to solve problems, settle disputes, acquire and organize new information, etc. To illustrate, tailors and carpenters measure differently, not surprisingly as they are measuring different things for different purposes. They must learn task-specific skills; but this does not mean that there will be no general skills which might transfer from carpentry to tailoring. Clearly tailoring and carpentry require very similar specific skills (measuring, cutting, fitting, joining) carried out in different mediums. But, in addition, there are general control factors involved in these and many other tasks. Asking one's self, "What is the first step?" applies equally well to data analysis, dressmaking, or shelf hanging. Similarly, self-interrogations such as, "Does this outcome make sense (satisfy, meet the goal set, etc.)?" is a self-monitoring operation necessary for the rejection of a negative probability result in a statistics problem, or a misplaced seam, or of a slanting shelf.

Such metacognitive skills may indeed be context bound to a degree: the master carpenter may not think to use them when faced with a syllogism, and the logician may not think to use them when hanging shelves. But, by their very nature, they have more likelihood of generalizing than content-specific skills. Interesting questions which we cannot answer on the basis of the extant data base are the extent to which such general skills do transfer across task domains, are or are not influenced by formal schooling, and can be taught directly.

We would, therefore, like to see examinations of the effect of schooling at the level of metacognitive skill. The difference is one of emphasis but, we think, not trivial. We realize that we are providing a restatement of the original problem—do educated (school experts) differ from uneducated (school novices) only in school-like situations or generally in their ability to deal with new problems? But the question whether expertise is context bound is an important one for theories of cognition in general and evaluations of the cognitive consequences of schooling in particular.

We argue that a refocusing of attention on metacognitive aspects of problem solving is desirable for three reasons: (1) there are "real-life" counterparts to such skills; (2) the skills are *by definition* transituational, and metacognitive activities are often implicated in learning to learn and transfer situations; and (3) one could make the case that inculcating such skills is the hidden curriculum of formal schooling. We will consider each point separately.

The skills of metacognition are essential prerequisites for efficient performance on a wide variety of tasks. Planning one's problem-solving strategies, monitoring the success of such attempts, checking the outcomes against criteria of efficiency, reality, and suitability are activities involved in many everyday reasoning situations, such as social interactions (settling arguments or disputes), economic planning (decisions to take one's wares to market), and navigational decisions, etc. (Cole et al. & Laboratory for Comparative Human Cognition 1978a). As such, this broad class of cognitive activities clearly has "everyday reasoning" counterparts and features prominently in school success. We argue that a refocusing of attention on the transfer of general skills is needed before one can conclude that schooling does or does not exert a significant influence on everyday reasoning.

A second reason for recommending the reemphasis is that the crucial question is one of transfer of training, the conditions promoting flexible, general use of the routines available to the problem solver. It has been argued extensively elsewhere that transfer of training requires metacognitive skills of prediction, monitoring, relatedness seeking, etc. (Brown & Campione 1978), and therefore we will not repeat this argument here.

Third, at a much more speculative level, we would like to argue that schools are, at least implicitly, attempting to inculcate just such general skills of flexible thinking that we include under the heading metacognition; for it is a common stricture that schools should teach children how to think rather than to deluge them with specific content which soon may become outdated. One way in which schooling may implicitly foster the ability to learn new information and solve novel problems is through instilling an awareness of whether information being presented is understood. If we "know we don't know," then this knowledge can in turn lead to self-questioning routines such as, "What do I know that might help me figure it out?" "What specifically do I understand?" "Where can I go to find out?" Termed "metacomprehension," the awareness of whether or not we understand something is a very general cognitive function with implications for almost every learning endeavor in which we might engage (Brown 1978).

In contrast to formal schooling, consider the process of training in a one-to-one situation, in which the teacher is an expert in whatever he is teaching and is more interested in directly transmitting the essential information or skills to the learner than in engaging the learner in a Socratic dialogue. In this situation, the expert monitors the learner's performance and can notice and correct any misunderstandings without there ever being any need for the learner to become aware that he has failed to understand. Such a learning situation corresponds very closely to the description given by Scribner and Cole (1973) of informal education and apprenticeship systems. It also bears a close parallel with Wertsch's (1977) description of problem solving by mother-child diads, in which the mother takes on the executive planning function for the child. In both cases the expert guides the novice's experiences in such a way that he does not have to plan for himself.

Formal schooling, at least in the later grades, is assumed to be quite different. Instruction is carried out in groups, outside of the applied context. The main vehicle for transmission of information is symbolic/verbal rather than observation/activity. In such situations, the students must learn to monitor their own comprehension, because the teacher in a large classroom cannot perform this function for them. To receive assistance the student must realize that he needs it and must know how to request it. Generally, after a few years of formal education, students are asked to acquire much of their information from books. Learning through reading makes it even more crucial that the student be able to monitor his comprehension, because there is no chance that a book will notice the student has failed to understand.

Learning to recognize and deal with one's own ignorance may lead to long-term benefits quite apart from the specific content being learned. Of paramount importance is whether schooling promotes this type of self-awareness and whether such skills transfer from school task to everyday reasoning.

It is unfortunate that comparisons between educated and uneducated populations have focused exclusively on school-type tasks, because, as a result of this limited focus, the conclusions which can be drawn are ambiguous at best; we cannot tell on the basis of current research whether to attribute the superior performances of schooled subjects to their having repeated experience with a class of tasks used in both the school and the experimental setting, or to the development of general, context-free problem-solving strategies.

Ideally, what is needed to clarify the ambiguity is an examination of situations which require schooled and unschooled participants to learn new tasks unrelated to the activities fostered by either schooling or the occupational pursuits of the unschooled adults. Then one could ask: Who will do better? In what ways will the groups differ? If the schooled populations were to do better and to display their superiority in the domain of planning and monitoring activities, we would be more justified in suspecting that formal schooling promotes general strategies for learning. It could be that metacognitive skills develop both within school settings *and* in a variety of everyday problem-solving domains. But the nature of the schooling process could be to make the general rules more explicit and bring them under conscious control so that they are used more readily, consistently, and in a wider range of situations. Education would serve also to join such activities to a whole new class of intellectual skills such as mathematics and literacy. An examination of the effects of formal schooling on the development and control of metacognitive skills has the potential for providing much useful information to reinforce the very clear picture to emerge from the Yucatan project.

Education and Social Policy

The Yucatan project has provided impressive evidence of the advantages of schooling, but Sharp et al. stress that great care should be exercised when interpreting their data because important social policy decisions may be influenced by these interpretations. The general argument is that if schooling promotes school skills and little else, then advocacy of universal education for Third World countries must be tempered with appropriate caution. This is clearly an important warning, but there are two levels on which this argument has been conducted by the Cole group. First, there is the question of whether schooling in general, and literacy in particular, exerts a fundamental influence on the course of cognitive growth and hence molds the very mode of thinking of its recipients. This is an extremely complex and important question for both theoretical and practical reasons and, as Sharp et al. point out, we are a long way from being able to answer it given the paucity of existing data.

At a more mundane level, however, there is quite convincing evidence that schooling, to the extent that it is successful, does promote school-relevant skills, and these include the nontrivial accomplishment of basic literacy and accounting skills. We would argue that the acquisition of basic literacy skills does indeed have important implications for performance in nonschool settings and that, quite apart from more general cognitive consequences of formal education, the acquisition of literacy, for literacy's sake, should not be deemphasized. Even if schooling does not produce the idealized general problem solver, to the extent that it does provide access to written materials and facility with clerical and bookkeeping skills it must have important consequences for the occupational opportunities of its recipients. Note, in the context, that it is when schools in our society are judged deficient at inculcating the same basic skills of functional literacy that widespread dissatisfaction is expressed. Thus, while we applaud the caution taken by the Cole group when discussing the potential impact of universal schooling (Cole 1978), we would like to emphasize a separation of the effects of the acquisition of literacy per se and the general cognitive sequelae of education.

In conclusion, we would like to emphasize the extremely important contributions of the Yucatan project. Because we have focused primarily on the interpretations that can be made concerning the general effects of schooling, we have deemphasized the actual accomplishments of this project in providing such a clear picture of the difference between the educated and uneducated, a picture that is cohesive, informative, and uncontaminated by selection artifacts. That we are free to speculate about potential interpretations is itself an important indicator of the strength of the data base provided by the Yucatan project. These studies provide invaluable information for psychologists concerned with intellectual development both within our society and in cross-cultural perspective.

References

Brown, A. L. Knowing when, where, and how to remember: a problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology.* Hillsdale, N.J.: Erlbaum, 1978.

Brown, A. L., & Campione, J. C. Permissible inferences from the outcome of training studies in cognitive development research. *Quarterly Newsletter of the Institute for Comparative Human Development*, 1978, **2**, 46–53.

Cole, M. How education affects the mind. *Human Nature*, April 1978.

Cole, M.; Hood, L.; & McDermott, H. Concepts of ecological validity: their differing implications for comparative cognitive research. *Quarterly Newsletter of the Institute for Comparative Human Development*, 1978, **2**, 34–37. (a)

Cole, M.; Hood, L.; & McDermott, H. Ecological niche picking: ecological invalidity as an axiom of experimental cognitive psychology. Unpublished manuscript, Rockefeller University, 1978. (b)

Cole, M., & Laboratory for Comparative Human Cognition. Cognition as a residual category in anthropology. *Annual Review of Anthropology*, 1978. (a)

Cole, M., & Laboratory for Comparative Human Cognition. What's cultural about cross-cultural cognitive psychology? *Annual Review of Psychology*, 1978. (b)

Flavell, J. H., & Wellman, H. M. Metamemory. In R. V. Kail, Jr., & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition.* Hillsdale, N.J.: Erlbaum, 1977.

Lave, J. Tailor-made experiments and evaluating the intellectual consequences of apprenticeship training. *Quarterly Newsletter of the Institute for Comparative Human Development*, 1977, **1**, 1–3.

Scribner, S., & Cole, M. The cognitive consequences of formal and informal education. *Science*, 1973, **182**, 553–559.

Wertsch, J. V. Adult-child interaction and the roots of metacognition. *Quarterly Newsletter of the Institute for Comparative Human Development*, 1978, **1**, 15–18.

[**Ann L. Brown** (Ph.D. 1967, University of London) is a professor of psychology at the University of Illinois at Urbana-Champaign. Her current research interests are in cognitive development and developmental disabilities. Two of her numerous publications are: "Knowing When, Where, and How to Remember: A Problem of Metacognition," in R. Glaser (Ed.), *Advances in Instructional Psychology*, vol. 1 (Hillsdale, N.J.: Erlbaum, 1978), and "Theories of Memory and Problems of Development: Activity, Growth, and

Knowledge," in L. Cernack & F. Craik (Eds.), *Levels of Processing in Memory* (Hillsdale, N.J.: Erlbaum, 1978).]

[**Lucia A. French** is currently completing her dissertation in developmental psychology at the University of Illinois at Urbana-Champaign. Her dissertation, on the long-term consequences of education, explores the cognitive skills of elderly women with different levels of schooling. Her future research plans include exploring the demands for reasoning and critical thinking which adults encounter in everyday settings and the skills needed to meet these demands. She has coauthored an article with Ann L. Brown and Judy De Loache, titled "IQ Tests in the Year 2000: Predicting the Adaptation of the Retarded Individual," which will appear in *Intelligence* (in press).]

* * *

REPLY BY MICHAEL COLE

The commentators on this *Monograph* have raised the key issues addressed in the body of the text itself in such plenitude that a full reply would probably amount to a text rivaling the original in length and hopefully exceeding it in clarity. In our comments we will try to eschew length and aim for clarity by choosing what we consider to be key points of conceptual unclarity among social scientists at large and ourselves in particular which we would like to see clarified in future research. An excellent starting point for this discussion is provided by Ginsburg when he writes, "It must be the case that uneducated subjects familiar with, say, navigation or car repair or hunting would require less task structure to perform the appropriate cognitive operations than would educated subjects unfamiliar with these areas" (p. 94). That speculation is central to our concern that the intellectual consequences of schooling, when assessed by tasks which take their structure from (or have a structure similar to) the tasks of the classroom, may be far more limited than we have previously been led to believe.

The side of the coin that we emphasized might be phrased in parallel with Ginsburg's speculation: "It must be the case that educated subjects familiar with, say, learning the names of all the states and their capitals or the periodic table would require less task structure to perform cognitive operations than would uneducated subjects unfamiliar with these areas."

From this perspective it is difficult to be as optimistic as Ginsburg seems to be about the strength of experimental approaches for drawing *general* conclusions about the "general human mind," or education's effect on that mythical entity. As we have argued in detail elsewhere (Cole, Hood, & McDermott 1978), such general conclusions must rest on demonstrations that models have generality in accounting for behavior across situations. That demonstration presupposes our ability to specify that the same task is occurring in widely differing environments (in particular, school and nonschool, "everyday" environments). We question our ability to make those crucial specifications regarding task structure. In our opinion, Ginsburg's suggestions that "experimental and Piagetian tasks are 'ecologically valid' in the sense that they refer to basic cognitive processes which must be used in the real world" (p. 96) misses the point. Even if we were to agree with Ginsburg's definition of ecological validity, we would not have dealt with the dilemma we posed.

In effect, we are arguing with Boas (1911/1965) *against Wundt*, when Boas says that "the existence of a mind absolutely independent of conditions of life is unthinkable. Experimental psychology, in its earlier stages, was sterile because it operated with the theory of the existence of an absolute mind, not subject to the environmental setting in which it lives" (p. 133).

Psychologists would like to believe that Boas's statement means no more than "environment affects the mind," *in general*. We are arguing that environments (here understood as microcontexts for thinking, such as those posed tests, vocabulary lists, navigation between islands and carburetors) bear the same relation to thinking processes in those environments. It is the implications of that position which are so difficult to live with because our methods for detailing with them are so paltry.

Brown and French, in their comments, seem to have accepted the seriousness of the dilemma in the terms in which we posed it. We agree with them that our data are both limited and dated by virtue of the techniques that we used. It is nice to know that a project begun in 1971 using techniques that were relevant then could be out of date 8 years later. But as our comments have made clear, we must be very careful in the way that we approach the problem of discovering "multicontext" cognitive tasks which Brown and French advocate.

A simple change in focus from "cognitive" to "metacognitive" tasks will not easily overcome the difficulties we have enumerated. The difficulty is raised by Brown and French when they point out that the master carpenter may not use metacognitive skills (What is the first step? Is that a reasonable answer?) when faced with a syllogism. The applications of metacognitive skills like applications of cognitive skills are dependent upon knowledge about the task. Thus, even when we change focus, we must do so while self-consciously sampling tasks "of the same sort" from domains of experience that are common to educated and uneducated alike, as well as problems which are more likely to reflect the special experiences of each group. Perhaps when this is done we will find ways to tap metacognitive skills that allow firmer generalizations about the effects of schooling. But the investigator who follows this route may be in for some surprises when he or she begins to examine everyday tasks for the extent to which metacognitive skills are applied in their solution. Brown and French imagine that questions such as "What is the first step?" are questions that the subject asks of oneself. Our observations (Cole et al. 1978) are that this question is more likely to be asked of someone else or provided, unasked for (in any obvious way), by others in the social environment. Moreover, feedback about the course problem solving and even a good deal of the "mental" work is likely to be done by others who are part of the scene and trying to get their work done. In other words, we may discover that the structure of non-school task environments, embedded as they are in larger scenes, are structured so differently that specifying "the same task" in widely different contexts will be very difficult. As emphasized in the Soviet literature on cognitive development (e.g., Istomina 1978; Vygotsky 1978), thinking is most often a means to some other end than an end in itself. Cognitive psychological experiments are exceptional in this regard, for "thinking" is both their objective and their content.

Our research not only supports this contention but suggests that, except when special care is taken to constrain individuals to do their thinking "in their heads," the mental work gets partialed out as an incidental feature of getting on with the *real* task (e.g., making certain that the crew on the little boat gets to the next island, or that the carburetor is repaired by 5 P.M.).

Thus, while we are in full agreement with the goals put forth by Brown and French, we are not sanguine about the ease of the enterprise on which they suggest we embark.

With all of these difficulties before us, we think that it would be worthwhile to suggest that future research consider bifurcating itself in a way not suggested in the previous discussions. All of the previous discussion has turned on the problem of specifying *cognitive processes* related to cognitive *development*. Our experience has led us to doubt the strength of the data for making general conclusions regarding either of these hypothetical entities on the basis of experimentally derived data. The alternative we suggest is that those interested in the cognitive consequences of schooling choose one of two goals, restricted so that the conclusions they draw will not wander too far from the data they collect. First, they can accept the limitations of their methods and pursue their search for cognitive process specialization with these limitations firmly in mind. Cognitive tasks are environments for telling us, in great detail, what people are doing. So long as we value conclusions about what people are doing in the kinds of environments they represent (e.g., because we want people who can run offices and construct microcomputers and figure out how we will be warming our houses in the year 2000), we can draw conclusions that are more or less safe about the activities that different people engage in within those environments (depending upon the skill with which we do our experiments and our accuracy in judging their representativeness). But we may not use those environments to talk about "the general human mind," except at the peril of talking nonsense.

On the other hand, if we want to assess the social consequences of schooling in terms of the new ways that people approach and solve problems, we might want to eschew microprocess conclusions and content ourselves with more general indicators of change. Do rural Yucatecans who have been to school more rapidly adopt advantageous strains of corn? Do they engage in better health care practices? Is their income as farmers raised? If the answers to those questions are affirmative, no link between individual mental development and economic development need be posited. We need no longer assume that the fact that people are economically underdeveloped or undereducated bears any close relation to their mental development. When we approach the problem using real-world activities as criteria of evaluation, we may begin to learn some very important things about the consequences of education that our cognitive tasks "underlook."

References

Boas, F. *The mind of primitive man.* New York. Free Press, 1965. (Originally published 1911.)

Cole, M.; Hood, L.; & McDermott, R. Ecological niche picking: ecological invalidity as an axiom of cognitive psychology (Working Paper). Laboratory of Comparative Human Cognition, Rockefeller University, 1978.

Istomina, Z. M. The development of voluntary memory in pre-school age children. In M. Cole (Ed.), *Soviet developmental psychology.* White Plains, N.Y.: Sharpe, 1978.

Vygotsky, L. S. *Mind in society.* M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.). Cambridge, Mass.: Harvard University Press, 1978.