Table 2. Proportion of errors in each of the five categories.

Category	Group 1	Group 2
i	0.20	0.14
ii	.12	.14
iii	.47	.30
iv	.53	.49
v	.25	.18

pairs. Consideration of the relative difficulty of the different types of pairs was restricted to those subjects who had at least one pair correct and at least one error. The proportions of errors for the different categories in Table 2 were therefore based on 17 subjects for group 1 and 19 subjects for group 2. A separate Friedman two-way analysis of variance by ranks was done for each group. Differences between categories were highly significant in each case. Chi square for group 1 was 320 and for group 2 it was 353. Significance at .001 in each case required a chi square of only 18.5.

As in the observations that led to this experiment, group 1 had more difficulty reversing common English sequences than other pairs; the proportions of items missed in categories iii and iv were far greater than in categories i and ii. The anomalous pairs in category v, however, were not much more difficult than pairs in categories i and ii. Thus, the fact that a pair forms a grammatical sequence seems to contribute relatively little to its difficulty if it is not a common English sequence. From a consideration of group 1 alone, it would appear that the absurdity of a pair when reversed does not contribute to its difficulty; pairs that were grammatical when reversed (category iv) seemed about as difficult as those that were not (category iii). As we shall see, this conclusion is unjustified when group 2 is also considered.

In group 2, as in group 1, greatest difficulty was encountered in categories iii and iv. The proportion of items missed in categories i and ii was much lower. As in group 1, anomalous pairs were almost as easy as those in categories i and ii.

The similarity in pattern of error in groups 1 and 2 suggests that the ability to separate items in a pair is of primary importance in the ability to reverse them. In addition, it should be recalled that the number of complete failures and complete successes was very similar in the two groups. Further evidence of the overlap of the abilities to separate items and to reverse them was obtained by testing the 23 failures on the other task 1 week later. Of the 13 children who failed in group 1, two were able to do the task of group 2. Of the ten children who failed in group 2, one was able to do the task of group 1. All of these measures indicate that the task set group 2 was slightly easier than that set group 1. This would be expected, since reversal of pairs is an additional operation after separation of the items in a pair.

Only in category iii was reversal far more difficult than separation of items. This is the category in which group 1 had to transform common English sequences into absurd pairs. Thus absurdity of the resultant pair also seems to be a factor in the difficulty of reversing the order of items.

The experiment was not designed to distinguish the relative difficulty of separating different English sequences into their component words. It would be expected, however, that difficulty would vary in different sequences, and in fact, categories iii and iv were not equivalent. Very small children are taught certain words singly, particularly object names. Even certain verbs and adjectives may be taught separately. On the other hand, certain words are probably rarely used in isolation; examples would be forms of the verb "to be," and personal pronouns. The confusion of the small child as to exactly what is a single word might be expected to be greatest with sequences made up of words that have not been used in isolation. The attempt to find pairs that were grammatical when reversed (category iv) led to exclusive use of such words, and this could account for the fact that these were more difficult to separate than pairs in category iii.

JANELLEN HUTTENLOCHER
Center for Cognitive Studies,
Harvard University,
Cambridge 38, Massachusetts

References and Notes

1. J. Church, *Language and the Discovery of Reality* (Random House, New York, 1963), p. 63.
2. H. Werner and E. Kaplan, *Monogr. Soc. Res. Child Develop.* **15**, 51 (1950).
3. S. M. Ervin and W. R. Miller, in *Child Psychology*, 62nd Yearbook, National Society for the Study of Education (Univ. of Chicago Press, Chicago, 1963).
4. The crucial suggestion to include this group was made by George A. Miller.
5. Supported in part by grant No. MH 05120-03 from the National Institutes of Health to Harvard University, Center for Cognitive Studies, and in part by grant No. MH 06626-02 from the National Institutes of Health.

19 November 1963

Student Performance in New High School Biology Programs

Abstract. *Data on the effectiveness of the Biological Sciences Curriculum Study course were obtained from several sources, including experimental use with 65,000 students. Any of the three versions of this biology course can be taught to average and above-average 10th-grade students. Achievement on the associated comprehensive test is more closely related to the ability and sex of the student, the salary of the teacher, the proportion of graduates of the school who go to college, the size of the class, and the adequacy of laboratory facilities than to the version of the course.*

The Biological Sciences Curriculum Study (BSCS) High School Biology Blue, Green, and Yellow Versions are parallel 1-year courses intended for use with average and above-average 10th-grade biology students in American schools. These versions were developed and written over a 4-year period by more than 100 outstanding research biologists and high school biology teachers organized into writing teams (*1*). Each version includes text, laboratory manual, and teachers' guide for a balanced first course in high school biology. While all versions include the same nine basic themes, each has different emphases on the various topics of biology. Preliminary editions of each of the versions were written in the summer of 1960 and were used experimentally as a complete course in 1960–61 by 118 teachers with 14,000 students; the materials used for making the texts, manuals, and guides were revised in the summer of 1962 and were tested by 350 teachers with 52,000 students in 1961–62, prior to writing a final edition.

The evaluation made in 1960–61 was in the nature of a "feasibility" study, to determine whether the material included the proper topics to be taught to high school biology students and whether these were useful for teaching basic biologic concepts. The evaluation was derived from reviews by specialists, reports from BSCS experimental students, visits of observers to schools, and performance of students on new BSCS tests. The evidence indicated that the content and grade of the materials were appropriate although considerable revision was needed (*2*).

In 1961–62 the same type of evalua-

Table 1. Correlations between BSCS comprehensive final and variables associated with ability and environment. ($N = 8965$)*

Variables	SCAT	Salary	Laboratory score	High school graduates to college (%)	Class size †	Comprehensive final
SCAT		.12	.01	.28	−.09	.72
Salary			.24	.12	.08	.17
Laboratory score				−.01	−.06	.07
Percentage to college					.00	.24
Class size †						−.09
Version ‡						.05
Mean	34.84	$6289	204	3.7 §	.59	26.22
S.D.	9.78	1491	53	1.1	.49	8.21

* Correlations of .03 or greater are statistically significant at the .01 level. † Class size as obtained from the questionnaires was dichotomized. Each teacher indicated whether the average size of his BSCS classes was 29 or less (coded as 0) or 30 or more (coded as 1). The correlations between class size and other variables are, therefore, point biserial coefficients. ‡ The correlations between version and comprehensive final is an eta, obtained by an analysis of variance. It is significant at the .01 level. § The percentage of students going to college was coded; 3.7 corresponds to approximately 40 percent.

tion was undertaken with a testing program made more rigorous by the use of a control group and a broader range of tests. Tests specific to each version were given quarterly for the experimental group. All groups were given a battery of common final examinations (the BSCS Comprehensive Final, a BSCS Impact Test, and a standardized conventional biology test—the Cooperative Biology Test, copyrighted in 1948 and still in fairly wide use). All students also took a general ability test, the School and College Ability Test (SCAT), Form 2B, parts 1 and 2.

The major findings on the 1961–62 evaluation were as follows (*3*):

1) Judged by achievement on the BSCS quarterly tests and on the end-of-year final examinations, BSCS students were able to master the BSCS biology materials and to achieve the desired skills to the satisfaction of the BSCS and the teachers using the materials. Average and above-average students did well in all versions.

2) Average scores on both the BSCS tests and the conventional Cooperative Biology Tests for boys were consistently higher than those for girls in both experimental and control groups.

3) Achievement of 10th-grade BSCS students on the BSCS Comprehensive examination was significantly higher than was that of 10th-grade students in the control group. However the achievement of the control group was significantly higher on the conventional Cooperative Biology Test than that of the BSCS students. On the BSCS Impact Test results were not definitive.

4) On all tests used, there was high correlation of achievement with general ability. An analysis of variables of student performance in the BSCS course as measured by the BSCS Comprehensive Final examination also indicated that successful performance was related to the salary of the teacher, the adequacy of the laboratory (as measured on the BSCS laboratory checklist), the size of the class (small), and the proportion of the school's graduates going to college. These four variables were more closely related to student achievement than was the version taught (Table 1).

5) For this sample population, no significant relationship was found between achievement and the type (rural, urban, or suburban) or size of the schools, the length of class period, the number of periods per week, the expenditure per pupil, and such characteristics of the teacher as age, years of experience, and number of undergraduate and graduate hours in biology.

6) Above-average 9th-grade students in situations with relatively good teacher preparation and good biology laboratories were able to handle BSCS biology materials. Data tend to substantiate reports of students and teachers that BSCS biology would probably not be suitable for the average 9th-grade student.

7) The variables of the ability and sex of the student, the salary of the teacher, the adequacy of the laboratory, the size of the class, and the proportion of school's graduates going on to college account for about three-quarters of the variance in student scores on the BSCS Comprehensive Final examination. Although the contribution of the variables of size of class and adequacy of the laboratory to predictability is small, nonetheless such relationships are not generally regarded as having been adequately demonstrated prior to this time. In this study, a number of extraneous sources of variation in the learning situation were rigorously controlled. This control may be the reason it has been possible to identify the association between student learning and certain environmental factors that had previously been thought, but not demonstrated, to exist. It is not implied that these are direct cause-effect relationships. Rather the identification of these factors may make it possible to describe a school climate in which effective learning occurs.

One-fourth of the variance in student achievement is still unaccounted for. It is still necessary to identify the causes of the variance as well as to ascertain whether the differences in student performance are short-term differences or whether they persist (*4*).

HULDA GROBMAN
Biological Sciences Curriculum Study, University of Colorado, Boulder

Notes

1. The BSCS was organized by the American Institute of Biological Sciences in January 1959, and has been supported primarily through National Science Foundation grants.
2. The BSCS tests were developed in cooperation with Educational Testing Service (ETS), which administered and scored them; the experimental design was suggested by ETS, which also selected the control group. Statistical processing was done by the Psychological Corporation, the University of Alabama, and the BSCS. The evaluation activities have involved a large number of persons including BSCS staff and committee members and staff of the two testing agencies; I am the BSCS staff consultant responsible for coordinating BSCS evaluation activities.
3. For comparisons at the 10th-grade level among BSCS subgroups and between BSCS students and the control group, adjustments were made so that groups are equated for ability as measured by SCAT.
4. The Psychological Corporation statistical report of the evaluation is on file with the BSCS. A more detailed general report of this study is included in *BSCS Newsletter 19*, in press.

8 November 1963