Br. J. educ. Psychol., 46, 164-173, 1976

THE DISTRIBUTION OF PIAGETIAN STAGES OF THINKING IN BRITISH MIDDLE AND SECONDARY SCHOOL CHILDREN

BY M. SHAYER, D. E. KÜCHEMANN AND H. WYLAM

(*Centre for Science Education, Chelsea College, London*)

SUMMARY. The proportion of children showing early and late concrete operational thinking, and early and late formal operational thinking was measured in a sample of 10,000 children between the ages of 9 and 14. The test instruments were a form of group test called Class Tasks, derived from the individual interview situations described by Piaget. These tests indicated that most children in early adolescence showed rapid development in concrete thinking, but that only one-fifth of the children showed the further development of formal operational thought. The representativeness of these findings was ensured by relating the distribution of Piagetian stages at each age-level to the norms of a standardised non-verbal reasoning test.

INTRODUCTION

As part of a study of the conceptual demands made in school mathematics and science curricula (CSMS project), it was thought desirable to know how the levels of thinking which Piaget describes are distributed in the school population. Do all children go through the stages described in *The Growth of Logical Thinking* (Inhelder and Piaget, 1958) at the ages mentioned there? The follow-up study of Lovell (1961) confirms the unity of Piaget's description, but was made on far too few children to show the population change year by year. Work by Tisher (1971) and Karplus (1975) suggests that the only economical way of testing the number of children required is to develop a test instrument which can be administered to the whole of a school class at a time.

The population initially chosen was that of the first three years of secondary education (11 to 14), but this age-range was extended to include 9 and 10-year-old children, because of the current interest in science for middle schools. In order to sample the whole range of thinking to be found in such a wide age-range it was necessary to develop measurement techniques from pre-conceptual thinking right through to late formal operational thinking. For this purpose three tasks were devised: the first covered pre-conceptual thinking (Stage 1) through to late concrete operational (2B); the second from early concrete (2A) to early formal (3A), and the third from late concrete (2B) to late formal (3B). The validity of each task was established by ensuring that each item had its counterpart in one of the original clinical interview situations described by Piaget and his co-workers.

DESCRIPTION OF TASKS

Task 1: This was based on *The Child's Conception of Space* (Piaget and Inhelder, 1956). Four situations are taken, and each may be scored at a number of levels characteristic of performance.

(*a*) An empty jam-jar is held up and each pupil is asked to draw it in cross-section, imagining it half-full of water. Then it is tilted to 45°, and the pupil asked to draw it as it would look if it were still half-full of water. Finally, it is held horizontally, and the same question asked.

(*b*) The children are asked to draw the outline of a mountain, and on its side draw a house and some trees.

(*c*) A jam-jar with a plumb-line hanging inside is provided for each child, who is asked to handle it. Then the same questions are asked as in (*a*).

(*d*) The child is asked to imagine that he is standing in the middle of a long straight road going away into the distance, and either side of it are rows of trees. He is asked to draw it as it would look.

Depending on the item scores an overall assessment on this task can range from pre-conceptual through to late concrete (2B), with an additional scoring of 2B+.

Task II: This was based on *The Child's Construction of Quantities* (Piaget and Inhelder, 1974), and was chosen as being particularly suitable for the range of measurement 2A to 3A. Unlike Task I, Task II is hierarchically organised, with each item scored right or wrong. There are 16 items, of which the first two are the classical water-pouring tests from Chapter 1 of *The Child's Conception of Number* (Piaget 1952) and the next is a substance-conservation question based on maize being 'popped' in front of the class. These three items test conservation of substance, an early concrete operational concept (2A). The next seven are all scored as late concrete (2B) items, and involve intuitive density and water-displacement concepts based on a block of plasticine being lowered into water in measuring cylinders, and then distorted in various ways. Then there is a 2B/3A item in which pupils are asked to hold a block of brass and a block of plasticine of the same dimensions, and asked how the amount of water they would displace would compare. Finally, there are three 3A items requiring an analytical concept of density for their solution.

Task III: This was based on Chapter 4 of *The Growth of Logical Thinking* (Inhelder and Piaget, 1958), also using Somerville's more detailed description of performances on the pendulum problem (Somerville, 1974). Task III has 14 scored items, again scored right or wrong, and each labelled 2B, 3A or 3B. On both this and Task II the overall assessment principle is to ask for success on at least two out of three of the items characteristic of a stage, or sub-stage. A simple pendulum is demonstrated, and the values which possible variables could take are dichotomised : long or short ; light or heavy ; gentle or hard push. Two readings are taken in front of the class of the number of swings the pendulum makes in 30 seconds; the values taken are such that from the two readings no clear deductions can be made. The children are then challenged to make deductions. Then follow four items scored at the 3A level, in which, given the first two readings, the children are asked to say what combinations they think they should take next in order to decide about the effect of length, weight, and push, respectively. Finally, two more readings are added to the first two, from which the effect of length and weight can be deduced, and also, with difficulty, that of push. The 3B items are obtained by asking the children to choose the right strategy by which to make these deductions.

VALIDITY OF THE TASKS

Since the criterion of validity chosen is the match of the task with the original Piagetian research, it seemed best to obtain a measure of it by comparing how the same children performed in the Class-Task situation, and in the original clinical interview from which the task items were taken. Professor Lovell, of Leeds University, kindly agreed to make a small-scale comparison in respect of

Tasks I and II; the CSMS team, with advice from Susan Somerville, made a larger-scale comparison on the Pendulum.

Task I. Seven children who had done the Class-Task were interviewed in their school by Professor Lovell, three-and-a-half months later, following the procedure outlined in *The Child's Conception of Space* (pp. 171, 173, 381-382). There were four interview items matching the four task items for each child, making 28 comparisons in all, together with a 'Points of View' item, which could be compared with the overall result on Task I. For these 35 comparisons a product-moment correlation of 0·85 was obtained.

Task II. The interview was based on *The Child's Construction of Quantities* (pp 4, 47-48, 119, 234, 258, 259). As it was rather lengthy, only four children were interviewed by Professor Lovell, about three months after they had done Task II. In three cases the Class-Task and interview item assessments were as close as in Task I, but the fourth case, a less able girl aged 13, shown as 2A or less by Task II, was given an overall assessment of 2B by interview, during which she appeared to have learned.

Task III. The interview and its assessment was based on Somerville's (1974) procedure. 54 boys and girls were interviewed six to eight weeks after they had done Task III. In this case the product-moment correlation was only 0·51. Task III had a 3B criterion based upon Piaget's analysis (Inhelder and Piaget, 1958, pp. 67-76): the interview had a lower criterion based on the children's responses, and this led to a systematic difference in assessment. Yet when the Friedman two-way analysis of variance was used to compare the results of the same children on Task II and Task III (Siegel, 1956, p. 166), although there were 22 cases of higher ranking by Task III than Task II, the probability that the two tasks were ranking the child alike was 0·8 ($\chi^2=0{\cdot}36$, $N=333$).

METHOD

Sample.

To obtain representative distributions of Piagetian stages in a British child population it was necessary to choose schools from a wide variety of areas, and to ensure that the sample of schools provided an appropriate proportion of children at different ability levels. Although the development and standardisation of the tasks used was done in London comprehensive schools, the one type of school absent from the survey was the large metropolitan comprehensive. 46 schools were used: four comprehensives in the Cardiff area, seven comprehensives in rural and city Gloucestershire, six small comprehensives in Hertfordshire; three comprehensives in rural and city Lincolnshire, and one comprehensive in each of Manchester, Staffordshire, Nottinghamshire and Southampton. 13 middle schools in the Leeds area were used for the 9- and 10-year-old children. To ensure adequate representation of the top 20 per cent of the ability range, and because of their interest as a separate population, nine selective schools were used: one each from Cambridge, Surrey and Nottinghamshire, and six in the London area. Each school was asked to ensure, as a minimum requirement, that every class in at least one year-group should be tested. The aim was to obtain for each year-group 1,200 children with complete data; to do this an initial sample of some 2,000 for each year-group was chosen, giving an overall total of about 10,000 children. The ability range of each school was measured by administering the Calvert Non-Verbal reasoning test (Calvert,

1958) to the whole of the 11+ year of the school (or to the 12+ year if that was not possible). (This test was chosen because it seemed to have, of the available tests of intelligence, most in common with what the Class Tasks were testing, and this is borne out by the correlations). Since this test has been standardised by the NFER on a substantial representative sample of British children to a mean of 100 and standard deviation of 15, comparison with this norm provided an indication of the representativeness of the sample used here.

Procedure for testing.

Although each task had full written instructions for administration and assessment, every school was visited by a member of the CSMS team for briefing on the tasks: where possible the task was demonstrated to the whole science department ; otherwise it was demonstrated to the head of science who then briefed his department. Testing was carried out in the academic year 1974-75. To avoid making an unreasonable demand on a school, only one task was administered in each term (in the autumn term, the Calvert Non-Verbal test was also given), and the team undertook, on each occasion, to keep within the bounds of a double-lesson (about 50 minutes). The class teacher demonstrated the problem and asked all the questions orally ; the questions were also printed on the pupil's response sheet. Pupils were allowed to ask questions to make sure they understood what they were being asked and the teacher was encouraged to make sure his pupils did understand, and to talk through the question with individual pupils if necessary. Wherever possible the answer could be made by drawing, or by ticking a right answer from a choice of possible ones. Finally, the teacher assessed the pupils' responses according to a schedule provided.

This test situation thus has some of the features of the interview, and some of the features of a psychometric test : the term ' Class-Task ' was coined to describe its difference from both interview and test. In this approach the teacher's skill, and his experience of his class, are used to make sure that the test situation is clearly understood. Accuracy of assessment was approached by a number of checks. For each class there was a 10 per cent check by a member of the CSMS team. If there were two or more differences in the assessment of an item in the response sheets of any three of the pupils, all of that class's response sheets were re-assessed. If it was clear from the pupils' responses that the test had been faultily administered or that the teacher had influenced the pupil's responses, that class's results were not used. Very few classes had to be omitted. Later all the assessments were recalculated from the item scores after item-analysis had shown that slight changes in the scoring-rules were needed.

RESULTS

Since there were only a few schools able to provide children from several years, each year-group was examined separately. For each school the Calvert mean standardised score, and the percentage of children at each Piagetian level assessed by a task was calculated. The regression line of task level percentage on Calvert mean was used to allow every measurement to contribute to the population estimate, and the level percentage corresponding to a Calvert mean of 100 was taken as representative. A check was made that this gave the same result as taking a group of schools with a combined Calvert mean of 100 and standard deviation of 15. The results for the three tasks are given in Tables 1, 2 and 3.

TABLE 1

PERCENTAGE OF CHILDREN AT DIFFERENT STAGES BY TASK I.

Child population		Stage of performance in Task I				
N	Mean age of children	1	2A	2A/2B	2B	2B+
1089	13 years 8 months	0·4	0·7	5·9	44·4	48·5
2209	12 years 8 months	1·5	3·8	16·0	45·7	33·0
2173	11 years 9 months	3·2	8·3	28·6	40·4	19·5
1132	10 years 8 months	8·2	12·4	33·0	36·6	9·8
831	9 years 8 months	11·0	17·0	36·2	28·2	7·7

TABLE 2

PERCENTAGE OF CHILDREN AT DIFFERENT STAGES BY TASK II.

Child population		Stage of performance in Task II					
N	Mean age	1	2A	2A/2B	2B	2B/3A	3A
1238	14 years 1 month	0·7	3·4	14·2	28·1	29·8	23·8
2094	13 years 1 month	3·4	5·8	21·0	29·3	26·3	14·2
1727	12 years 1 month	4·2	7·9	23·6	30·2	25·8	8·3
1157	11 years 1 month	7·3	13·8	36·0	27·2	10·1	5·6
765	10 years 0 months	11·6	16·9	36·8	25·2	7·4	2·2

TABLE 3

PERCENTAGE OF CHILDREN AT DIFFERENT STAGES BY TASK III.

Child population		Stage of performance in Task III				
N	Mean age	2B−	2B	2B/3A	3A	3B
1305	14 years 4 months	29·4	25·2	23·6	14·8	7·1
1543	13 years 4 months	35·3	23·8	23·1	14·1	3·8
1292	12 years 3 months	40·1	20·5	25·4	12·6	1·5

TABLE 4

PERCENTAGE OF SELECTIVE SCHOOL CHILDREN AT DIFFERENT STAGES.*

Child Population		Stage of performance in Task II				
N	Mean Age	2A/2B	2B	2B/3A	3A	3B
249 (159)	14 years 2 months	1·2	6·4 (0·6)	28·5 (13·2)	63·8 (86·2)	— —
150 (146)	13 years 2 months	2·0	8·0 (2·1)	34·7 (11·0)	55·3 (87·0)	—
430 (217)	12 years 2 months	9·0 (0·5)	18·3 (12·0)	39·0 * (31·3)	33·6 (56·2)	— —
		Stage of performance in Task III				
146 (135)	14 years 3 months	5·2	8·8 (2·9)	21·8 (11·6)	40·9 (27·3)	23·3 (58·1)
134 (157)	13 years 3 months	9·2 (1·0)	14·7 (2·0)	32·6 (16·8)	30·4 (34·2)	13·0 (46·0)
321 (266)	12 years 4 months	8·1 (3·4)	15·9 (7·5)	36·4 (28·2)	24·0 (34·2)	15·6 (26·7)

* Details for the ' super-selective ' sample are shown in parentheses.

In Table 4 there are two sets of results for the selective school population. Three of the schools are known nationally, and their pupils figure in the Oxbridge scholarship tables: their results are amalgamated separately, and given in parentheses. It is estimated that the population of these schools represents the top 8 per cent of the ability range. The other six schools were good grammar schools taking approximately the top 20 per cent of the ability range and their amalgamated figures are quoted as representative of the average English grammar school.

TABLE 5

CORRELATIONS BETWEEN TASKS AND CALVERT TEST, BETWEEN DIFFERENT TASKS, AND RELIABILITY ESTIMATES.

	Calvert Test	Task II	Reliabilities
Task I	0·60	0·41	0·82
Task II	0·55	—	0·80
Task III	0·63	0·57	0·86

Correlation coefficients and reliabilities are given in Table 5. The Task II and III reliabilities were calculated by KR-20, and for Task I by Hoyt's analysis of variance (Winer, 1970, pp. 124-128). Both the correlations and reliabilities were calculated on samples of between 300 and 400, and the age variable was largely eliminated by using only one year-group for each.

DISCUSSION

In order to show the overall pattern of development in the child population studied the results are plotted in Figures 1, 2 and 3, in the form of cumulative percentages. A point showing 50 per cent of the population of that age as 2B should be read as 50 per cent at that level or above. In the population as a whole the development of early formal thinking seems to be confined to the top 20 per cent of the year-group; when the top 20 per cent are looked at separately, most of their development has already taken place by age 14½. Thus, any rise in the proportion of young people beyond the age of 14 showing early formal performance must come from that part of the population which lies below this top group. Until the next two year-groups in the secondary school have been studied it will not be possible to read these curves unequivocally, but many teachers with comprehensive school experience would be surprised if development did not continue, at least for that part of the population between the 80th and the 50th percentile of ability, well up to 17 years of age. It is intended to compare these results with analyses of the apparent conceptual demand of science and mathematics curricula in an attempt to throw light on some of the learning problems of secondary and middle school children.

FIG 1

PROPORTION OF CHILDREN AT DIFFERENT PIAGETIAN STAGES IN A REPRESENTATIVE BRITISH CHILD POPULATION.

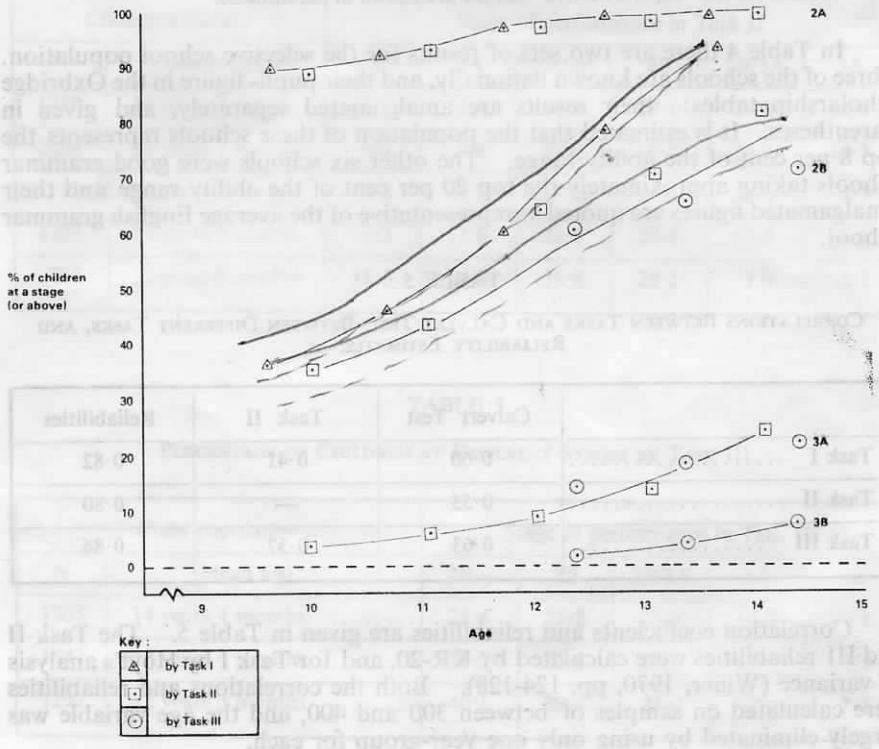


FIGURE 2

SELECTIVE SCHOOL POPULATION (TOP 20%).

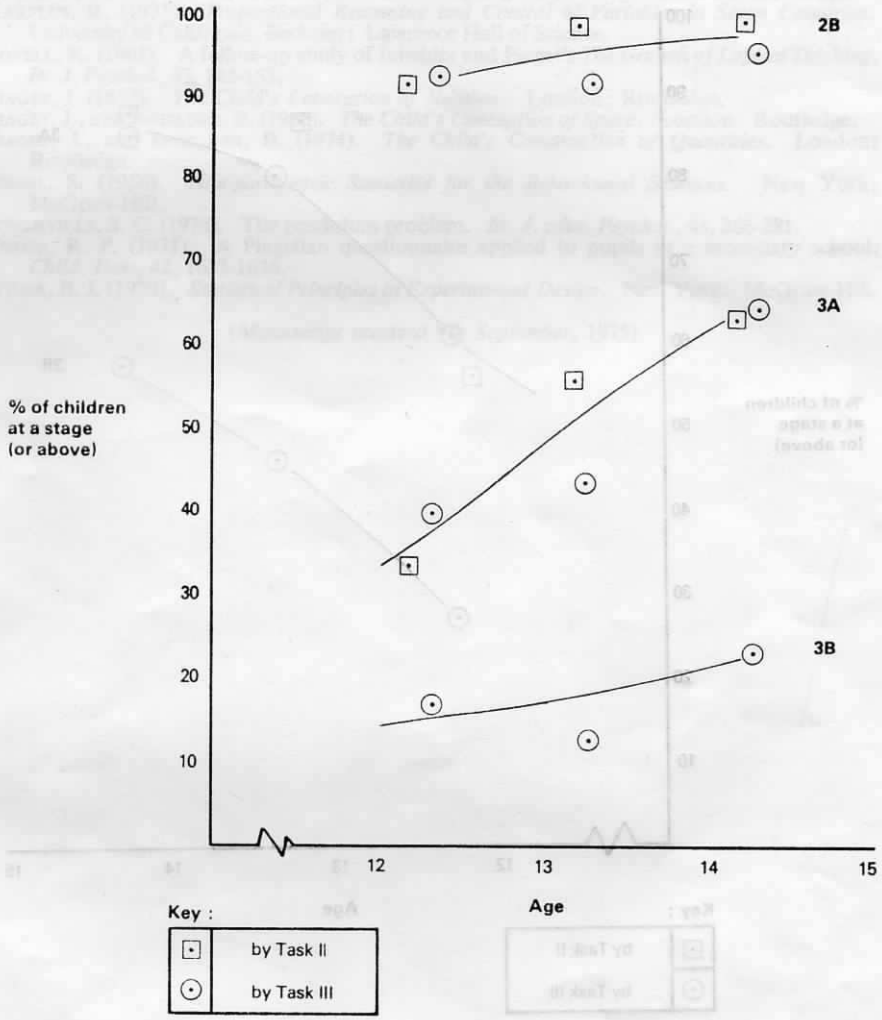
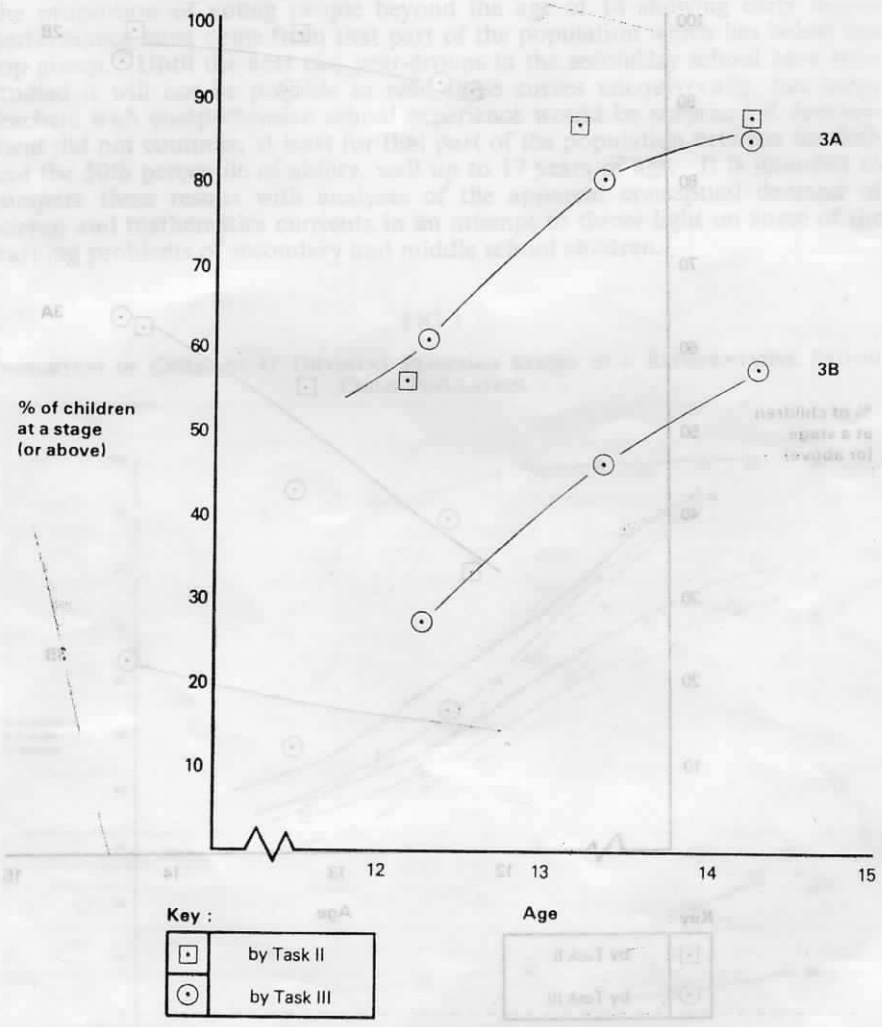


FIGURE 3

SUPERSELECTIVE SCHOOL POPULATION (TOP 8%).



ACKNOWLEDGMENTS—The work reported in this paper forms part of the research programme ' Concepts in Secondary Mathematics and Science ' (CSMS) funded by the SSRC at Chelsea College.

REFERENCES

CALVERT, B. (1958). *Non-Verbal Test DH.* Slough: NFER.
INHELDER, B., and PIAGET, J. (1958). *The Growth of Logical Thinking.* London: Routledge.
KARPLUS, R. (1975). *Proportional Reasoning and Control of Variables in Seven Countries.* University of California, Berkeley: Lawrence Hall of Science.
LOVELL, K. (1961). A follow-up study of Inhelder and Piaget's *The Growth of Logical Thinking. Br. J. Psychol.*, 52, 143-153.
PIAGET, J. (1952). *The Child's Conception of Number.* London: Routledge.
PIAGET, J., and INHELDER, B. (1956). *The Child's Conception of Space.* London: Routledge.
PIAGET, J., and INHELDER, B. (1974). *The Child's Construction of Quantities.* London: Routledge.
SIEGEL, S. (1956). *Non-parametric Statistics for the Behavioural Sciences.* New York: McGraw-Hill.
SOMERVILLE, S. C. (1974). The pendulum problem. *Br. J. educ. Psychol.*, 44, 266-281.
TISHER, R. P. (1971). A Piagetian questionnaire applied to pupils in a secondary school. *Child. Dev.*, 42, 1633-1636.
WINER, B. J. (1970). *Statistical Principles in Experimental Design.* New York: McGraw-Hill.

(*Manuscript receivcd 9th September*, 1975)