DOCUMENT RESUME

ED 051 583 EA 003 590

AUTHOR	Sumner, G. C.
TITLE	Project R-3 Allocation of Students Among Groups.
REPORT NO	P-4584
PUB DATE	Feb 71
NOTE	10p.; Paper presented at Association of California Administrators in Compensatory Education Meeting. (Oakland, California, March 8, 1971)
AVAILABLE FROM	The Reports Department, The Rand Corporation, 1700 Main, Santa Monica, California 95001. (#P-4584, $1.00)
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	*Academic Achievement, *Educational Experiments, Grade 7, *Heterogeneous Grouping, Rating Scales, Reading Ability, *Research Methodology
IDENTIFIERS	Mathematical Ability

ABSTRACT

This report describes an objective method for allocating San Jose students to groups so that each group would include the same representation of students according to two quantitative measures of scholastic achievement. The conditions provided that each group be internally heterogeneous, but that intergroup differences be small, so that each would represent a separate replication of the same experiment. The allocation was to mirror uniformly the central tendency and variability of the overall student population. This method could apply to a more flexible set of initial conditions: it could accommodate another variable, be equally useful for ensuring representation across socioeconomic or cultural variables, and obtain proportionality with the racial-ethnic mix of the student population considered. (Author)

PROJECT R-3 ALLOCATION OF STUDENTS AMONG GROUPS

G. C. Summer

February 1971

P-4584

PREFACE

The allocation technique described in this paper was developed as part of Rand's study of the San Jose Unified School District's Project R-3, a demonstration program conducted under the auspices of the California State Department of Compensatory Education as provided for in Assembly Bill 938. The technique provided the means for establishing classroom heterogeneity among the students in the project. Although the description is specific to Project R-3, the method generalizes to a rather flexible range of allocation requirements; this flexibility is also discussed.

PROJECT R-3 ALLOCATION OF STUDENTS AMONG GROUPS

G. C. Sumner*

The Rand Corporation, Santa Monica, California

Before the semester was underway, it was necessary to divide the seventh graders into class-size groups. The groups were to be internally heterogeneous with respect to sex and scholastic abilities, but between-group differences were to be small; thus, there would be no semblance of tracking and each group could be regarded, to some extent, as a separate replication of the same experiment. The guiding strategy was to establish an objectively reproducible selection procedure, a method free of intentional or unintentional bias.

There were to be 12 groups of equal size; each group would have proportional representation by sex and each would represent a full range of reading and arithmetic abilities as manifested in raw scores achieved on the CAT, which was administered in January.

At the start of the semester, February 2, 253 students were enrolled (136 boys and 117 girls). Five students had not taken all of the CAT tests, but were assigned proxy scores.† The mean and median

---

*Any views expressed in this paper are those of the author. They should not be interpreted as reflecting the views of The Rand Corporation or the official opinion or policy of any of its governmental or private research sponsors. Papers are reproduced by The Rand Corporation as a courtesy to members of its staff.

This paper was presented to the Association of California Administrators in Compensatory Education held at the Oakland Hilton Hotel, 8 March 1971.

†Proxies for three students were extrapolated from relative placements in the Title I testing administered the previous October:

$$ {}_i\tilde{X}_J = \bar{X}_J + S_J\left(\frac{{}_iX_o - \bar{X}_o}{S_o}\right), $$

where ${}_i\tilde{X}_J$ = proxy for the i*th* student,

$\bar{X}_J$ = mean score of the January testing,

S_J = standard deviation of the January testing,

$\bar{X}_o$ = mean score of the October testing,

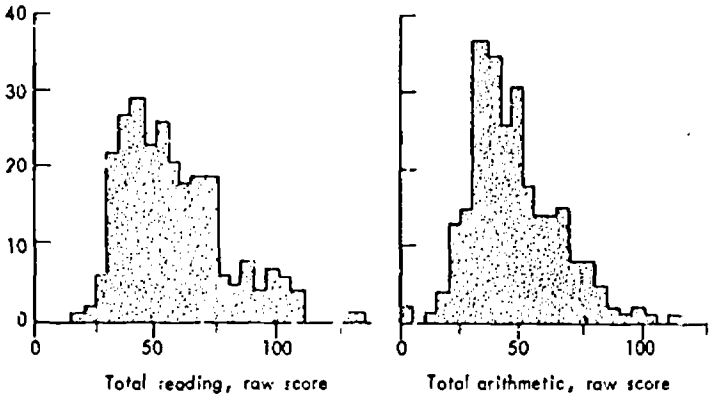
S_o = standard deviation of the October testing,

${}_iX_o$ = score received by the i*th* student on the October test.

Since the previous test included only Title I students, who presumably

raw scores were:

	Mean	Median
Total reading	57.52	53
Total arithmetic	47.26	44



We now describe the scheme used to allocate 252 students into 12 groups of 21, such that each group contains 3 students from each of 7 reading-ability rankings (reading septiles) and from each of 7 arithmetic-ability rankings (arithmetic septiles).† The remaining student was to be arbitrarily assigned to one of the groups.

The students were first ordered according to the reading raw scores. The top 36 students were assigned to the first septile, the

---

would score lower on the average, the respective assignments probably are positively biased.

Because the other two students were not included in the October testing, the January test data corresponding to the nonmissing score were substituted. For example, if the reading score was missing for the ith student:

$$i\bar{X}_{read} = \bar{X}_{read} + S_{read}\left(\frac{i X_{math} - \bar{X}_{math}}{S_{math}}\right)$$

†The choice of 7 ability rankings (rather than, say, 5 or 10 rankings) was primarily a matter of convenience. For the problem at hand, the number of rankings could have been as small as 2 or as large as 21. Smaller numbers would provide less heterogeneity and larger numbers would allow less freedom for allocating the sexes. The "sensible range" for this case would therefore lie between 5 and 12. Since the allocation procedure is more straightforward if the number of rankings is an exact divisor of the group size, the likely candidate for this case is 7 rankings.

next 36 students were assigned to the second septile, and so on down to the last (seventh) septile. Each student was assigned a number corresponding to his reading septile. Next, the students were similarly divided into arithmetic septiles, and assigned corresponding identifiers. Using the two-digit identifier thus assigned, the joint distribution of students across the two ability measures was charted on a 7 by 7 matrix:

Arithmetic Septiles

Reading Septiles	1	2	3	4	5	6	7	
1	21	8	4	2	0	1	0	36
2	7	11	11	3	2	1	1	36
3	5	9	4	7	4	5	2	36
4	2	5	5	12	3	6	3	36
5	1	1	8	8	3	6	9	36
6	0	2	1	3	16	9	5	36
7	0	0	3	1	8	8	16	36
	36	36	36	36	36	36	36	252

The numbers in the cells indicate how many students were in the respective categories; numbers on the margins provide row (reading septile) and column (arithmetic septile) totals.

A distribution pattern for each of the 12 groups was designated (in sequence) by choosing 3 non-zero cells from each row and from each column of the matrix.

After each pattern was designated, the cell numbers were decremented by the number of selections from the respective cells and the marginal totals were adjusted to reflect the new sums. When designating any particular pattern, it was permissible to select the same cell more than once; the selection rules were:

1. If the cell number was zero, the cell could not be chosen;
2. If the cell number was non-zero but less than one-third the margin total it could be chosen once;
3. If the number was one-third the margin total, the cell had to be chosen;

4. If the number was between one-third and two-thirds of the margin total, the cell had to be chosen once but could also be chosen twice;
5. If the number was two-thirds the margin total, the cell had to be chosen twice.

These rules guaranteed that pattern designation would proceed in such a manner that the 12 patterns could be exactly accommodated by the 252 students. Furthermore, the rules precluded the necessity to select a cell three times and minimized the need for double selections. Within these restrictions, care was taken to distribute each pattern fairly evenly over the matrix. The selection of the first two and last two patterns is indicated by the asterisks in the illustration below:

Pattern 1

21*	8*	4	2	0	1	0	36
7*	11	11*	3	2	1	1	36
5	9	4	7*	4	5*	2	36
2	5	5*	12*	3*	6	3	36
1	1	8*	8*	3	6	9*	36
0	2	1	3	16*	9*	5*	36
0	0	3	1	8*	8*	16*	36
36	36	36	36	36	36	36	252

Pattern 2

10*	7*	4*	2	0	1	0	33
6*	10*	10*	3	2	1	1	33
5*	8*	4	6*	4	4	2	33
2	5	4	11*	2*	6*	3	33
1	1	7*	7*	3	6	8*	33
0	2	1	3	15*	8*	4*	33
0	0	3	1	7*	7*	15*	33
33	33	33	33	33	33	33	231

Pattern 11

4*	1	1*	0	0	0	0	6
1*	1	1*	1	0	1*	1	6
0	1*	1	2*	1	1*	0	6
1	1*	1	2*	0	1*	0	6
0	1	1	1*	1*	1	1*	6
0	1*	0	0	4*	1	0	6
0	0	1*	0	0	1	4*	6
6	6	6	6	6	6	6	42

Pattern 12

2*	1*	0	0	0	0	0	3
0	1*	0	1*	0	0	1*	3
0	0	1*	1*	1*	0	0	3
1*	0	1*	1*	0	0	0	3
0	1*	1*	0	0	1*	0	3
0	0	0	0	2*	1*	0	3
0	0	0	0	0	1*	2*	3
3	3	3	3	3	3	3	21

The next task was to match students to the patterns designated. The patterns were taken one at a time and students with identifiers matching the chosen cells were selected. This selection was arbitrary (unsystematic) except that the ratio of boys to girls was either 11 to 10 or 12 to 9 and both sexes were distributed fairly evenly over the pattern. The extra student (number 253) was arbitrarily assigned to the ninth group; his test scores would have placed him in cell 6.6.

Finally, the 12 groups were shuffled and assigned the labels used in the R-3 program (A: 1, 2, 3, 4. B: 1, 2, 3, 4. C: 1, 2, 3, 4). The 12 groups are illustrated, along with the R-3 labels, in Fig. 1. in the order of designation. Boys are indicated by Bs, girls by Gs.

It may be of interest to examine how well this procedure performed its task of allocating students into groups that uniformly mirror the central tendency and variability of the overall student population. Table 1 provides comparisons with respect to reading and arithmetic scores. The groups are identified by their R-3 codes. Means and standard deviations are given for each of the 12 groups, for each of the 3 lettered groupings (A, B, and C), and for the overall population.

It is also interesting to compare the variability of the group means with that which might have occurred had the students been allocated by simple random methods. The standard deviation of the means of the groups is computed as

$$\sqrt{\frac{\sum_{g=1}^{12} (\bar{X}_g - \bar{X}_o)^2}{12}} \quad \begin{cases} = 1.6 \text{ (reading)} \\ = 1.6 \text{ (arithmetic)} \end{cases}$$

where $\bar{X}_o$ is the mean for all 253 students. Had groups been allocated at random, one could have expected the standard deviations to be in the neighborhood of

$$\frac{S_o}{\sqrt{12}} \begin{cases} = 4.6 \text{ (reading)} \\ = 3.9 \text{ (arithmetic} \end{cases}$$

where S_o is the standard deviation of the overall.

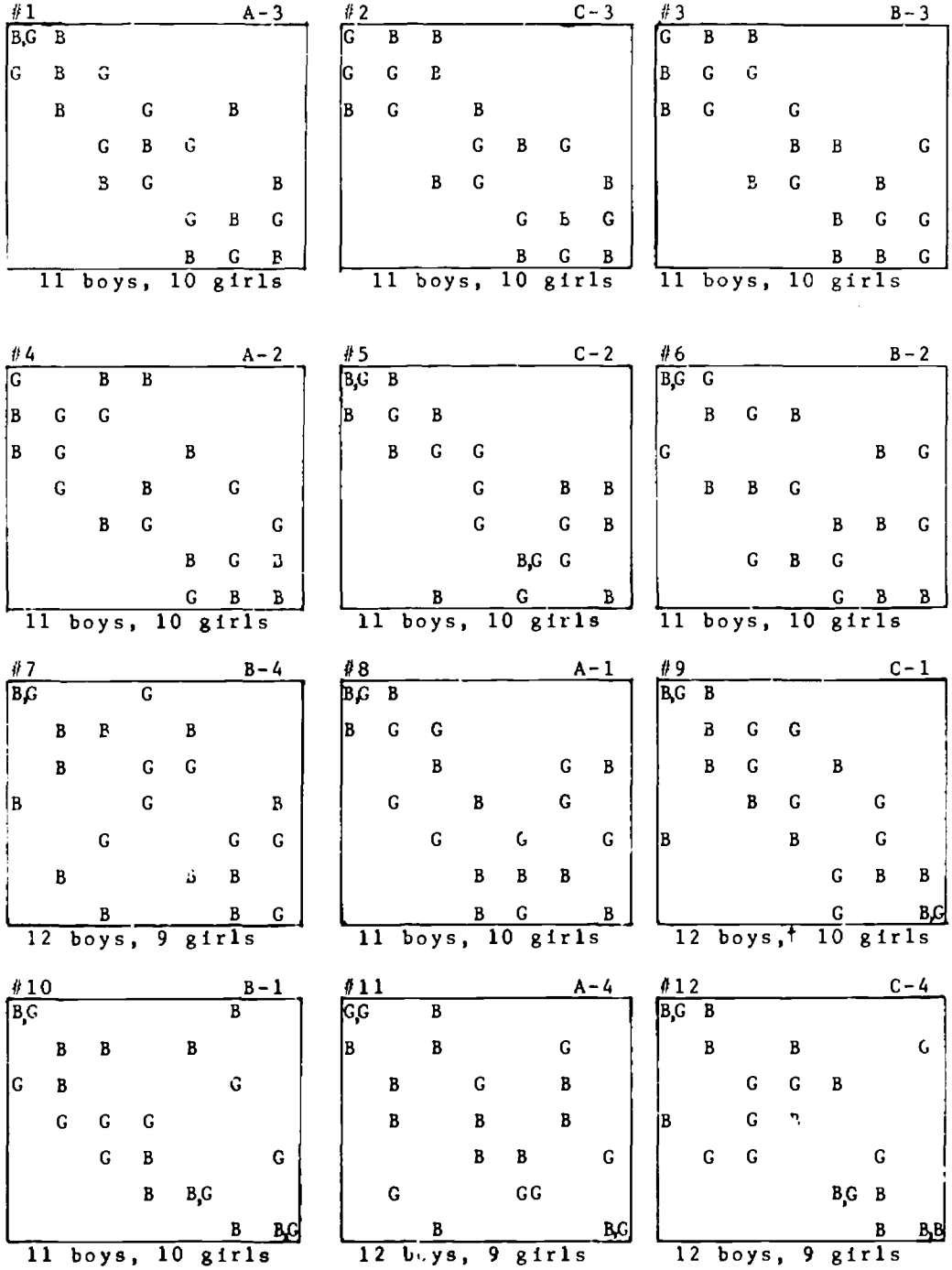


Fig. 1--Allocation of students by septile and sex among groups

†Includes the extra student.

Table 1

GROUP MEANS AND STANDARD DEVIATIONS

Group	Reading		Arithmetic		n
	Mean	Std. Dev.	Mean	Std. Dev.	
A-1	55.52	18.77	46.95	20.12	21
A-2	56.71	19.63	47.76	16.95	21
A-3	61.38	24.99	48.86	18.50	21
A-4	55.62	16.74	46.33	16.49	21
A total	57.31	20.41	47.48	18.10	84
B-1	56.29	21.66	47.29	17.79	21
B-2	57.43	22.80	47.29	15.87	21
B-3	58.86	21.64	47.19	15.74	21
B-4	56.71	19.85	47.81	18.01	21
B total	57.32	21.54	47.39	16.89	84
C-1	56.27	19.78	44.32	18.26	22
C-2	59.29	24.42	50.86	24.20	21
C-3	59.76	20.64	46.29	16.41	21
C-4	56.48	18.79	46.33	13.91	21
C total	57.93	21.06	46.92	18.79	85
Overall	57.52	21.01	47.26	17.93	253

The sex ratio in the larger group was 136 boys to 117 girls, or 11.33/9.75. Eight of the smaller groups were assigned 11/10 ratios, three had 12/9, and one had 12/10. Short of partitioning some youngsters, there is no allocation with ratios more uniformly near that of the overall ratio (136/117).

The preceding pages describe an objective method for allocating students into groups so that each group includes the same representation of two quantitative measures of scholastic achievement. The method can easily accommodate a more flexible set of initial conditions:

1 With more effort, a third controlling variable (i.e., ranking criterion) could have been formally incorporated into the procedure. For example, students could have been ranked according to their score on the language section of the pre-test. The two-dimensional matrix, p. 3, would then become a three-dimensional matrix characterized as having seven rows (reading septiles), seven columns (arithmetic septiles), and seven files

(language septiles). Then, the allocation to each of the twelve groups would be such that three students are included from each file as well as from each row and column. Even more controls may be added, but with increasing difficulty.

2. The control variables need not be quantitative; the method would be equally useful for insuring representation across socioeconomic or cultural variables. For example, in some compensatory education programs it might be desirable to allocate students along a three-way ranking scheme using reading pretest scores for one ranking criterion, racial or ethnic characteristics for the second, and a subjective assessment of English-speaking ability for the third.
3. Although it is administratively convenient for the control variables to split the student population into rankings of equal size, this is not necessary because only the number of students selected from each ranking need be proportional. Thus, in the example above, student allocation to groups should reflect the relative sizes of the racial-ethnic rankings. (This obtains proportionality with the racial-ethnic mix of the population being considered.)