simulation & games

"Games and Teams:
A Winning Combination"
by
Keith J. Edwards,
David L. DeVries,
and
John P. Snyder
is reprinted from
Simulation and Games,
Volume 3, Number 3,
(September 1972) pp. 247-269,
by permission of the Publisher,
Sage Publications Inc.

GAMES AND TEAMS

A Winning Combination

KEITH J. EDWARDS, DAVID L. DeVRIES, and JOHN P. SNYDER
Center for Social Organization of Schools, Johns Hopkins University

Nonsimulation games have been in classroom use for some time and are viewed by many as an effective teaching medium. Empirical support for this contention, however, is minimal.

Competition among student teams has also been advocated (Coleman, 1959; Bronfenbrenner, 1970; Spilerman, 1971) as an effective teaching device in the classroom, particularly when combined with the playing of nonsimulation games (Allen et al., 1970).

The present study investigated the effects on student achievement of the combined use of nonsimulation games and student teams in mathematics. The nonsimulation game employed was EQUATIONS (Allen, 1969), a game of "creative mathematics."[1]

NONSIMULATION GAMES

Most empirical studies of the effectiveness of instructional games as teaching tools deal with simulation games, and most of the results suggest that simulation games are effective for changing student attitudes (e.g., Boocock and Schild, 1968; Lee and O'Leary, 1971; Edwards, 1971a). However, the results of the studies examining the effects on student achievement are less clear (for reviews of research, see Cherryholmes, 1966; Livingston and Stoll, 1972). Because of basic structural differences between simulation games and nonsimulation games, it is questionable whether the results of these studies can be extrapolated to the game used in the present study.

Allen et al. (1970) have reported a significant increase (20.9 points) in nonverbal IQ scores for students playing the nonsimulation game WFF N' PROOF (a game of logic). The students in the WFF N' PROOF treatment played the game four hours a day, five days a week for three weeks, while a comparable group of control students received traditional instruction in mathematics. Although the demonstrated gain is impressive, using games in such massive doses is impractical for most educational settings. The present study involved a more traditional classroom setting in which the game of EQUATIONS was used to complement traditional instruction, not to replace it.

One reason why the structure of EQUATIONS should improve mathematics achievement is that mathematics skills are necessary for winning. The more math a player knows, the more he will win. Schild (1968: 151) points out that for a game to produce learning it should be constructed so that the skills, insight, or facts to be learned are clearly needed by the players in order to succeed in the game.

A second feature of EQUATIONS that should produce increased learning is the challenge structure of the game. In brief, a player wins by either correctly challenging another player's mistake or by being incorrectly challenged. Challenges are resolved when a player explains to his competitors why the proposed mathematical solutions are or are not correct. This structure creates a series of potential student peer-tutoring situations. In the few cases where no player can identify a correct solution, the teacher acts as a consultant.[2]

A third feature of EQUATIONS that facilitates learning is the tournament used to structure competition. Students of approximately equal mathematics ability compete individually in small groups of two to three players. Tables are arranged in a hierarchy based on ability, and a "bumping" procedure is used between rounds of the tournament to correct placement errors and to adjust for differential rates of student learning. The winners are moved up one table and the losers moved down one table in the hierarchy. Thus the probability of a given student receiving positive reinforcement (i.e., winning a given game) is approximately one-third and is the same for all students. In addition, the reinforcement is more effective because it immediately follows the appropriate student behaviors.[3]

STUDENT TEAMS

Adding team competition to the EQUATIONS game enlists some additional classroom forces that should make the game even more effective. As indicated above, the game allows for student peer-tutoring, but it does not encourage such tutoring. In fact, because each student wins only if he outperforms his fellow players, it is a disadvantage to share information with them. Adding teams to the game structure allows the individual players to share information with their teammates. Sharing of this kind is to the student's advantage because it helps him and his teammates improve and thus increase his team's score during subsequent games. Empirical support for the assertion that student teams result in increased peer tutoring is abundant (e.g., Wodarski et al., 1971; Hamblin et al., 1971).

AUTHORS' NOTE: *The authors acknowledge the contribution of Sylvia Hall, math teacher, and Robert Freeman, department head, Northern Parkway Junior High School, for making this research possible. We are indebted to the following people for their assistance: Gail Fennessey and Alyce Nafziger for administering the tests and printing the newsletters; Karen Jaworski and Linda Ferraris for typing the manuscript; Layman Allen for helpful suggestions on using* EQUATIONS; *and James Coleman, Samuel Livingston, and Steven Kidder for helpful comments on earlier drafts of the paper.*

This research was a cooperative project between the Academic Games Program and the School Organization Program of the Center for Social Organization of Schools, a Research and Development Center supported by grants from the U.S. Office of Education.

Farran (1968) systematically compared the effect of individual versus team competition in three simulation games. Farran's results revealed greater learning by students in the individual competition. However, three features of her experimental manipulation of teams must be noted in interpreting the results. First, she created fifteen-member teams, a number which has been shown by small group research (e.g., Slater, 1965) to be much too large for participation by the majority of the team members. Also, Farran's game task involved a group effort in which a few individual students could dominate. Finally, the feedback on game performance was given only the group level for the team competition condition, thus preventing any student on the team from assessing his own individual past performance. Small group research has shown that both individual accountability and feedback must be present in a team situation in order for teams to be effective (compare Glaser and Klaus, 1966). In light of these findings, the present study formed small (four-member) teams, and the game involved individual student performance and feedback.

STUDENT ABILITY

Advocates of simulation games have argued that games reduce the differential learning rates noted for low- versus average- and high-ability students (e.g., Boocock and Schild, 1968; Coleman, 1972). That is, game playing is said to result in comparable levels of learning by students of all ability levels. Empirical evidence supports this contention when the criterion is learning strategies for winning the game or actual amount of winning (Fletcher, 1971; Braskamp and Hodgetts, 1971; Edwards, 1971b). However, playing a nonsimulation game such as EQUATIONS may especially facilitate the learning by low-ability students of more general academic skills. The hierarchical tournament structure increases the probability of success most for the low-ability students. Also, playing a game should be most appealing to low-ability students.

In the present study, we hypothesized that playing EQUATIONS, using team-competition would have a positive effect on student achievement in mathematics. We also hypothesized a significant games-teams by student-ability interaction, predicting a greater positive experimental effect on the low-ability students.

METHODS

SUBJECTS

The subjects were 96 seventh-grade students from a large urban junior high school. The students were in four general mathematics classes. Two of the classes contained primarily students of average math ability and two contained mostly students of low math ability. The school had no formal tracking procedure, so there was considerable variance on math ability within each class.

DESIGN

The experiment was a 2 x 2 x 2 repeated measures analysis of variance (ANOVA) design, with the three factors being: (a) treatment (games-teams versus traditional), (b) math ability (low versus average), and (c) time (pretest and posttest). The experiment was conducted for a nine-week period. All four math classes were taught by the same teacher (a female in her first year of teaching). Two of the classes (one average and one low-ability) were assigned to the games-teams (experimental) treatment, while the remaining two were assigned to the traditional (control) treatment. The two experimental classes met during periods one and two, and the two control classes met during periods four and six. Although the time period factor is confounded in the design, its impact on the students' reactions to the treatments was considered by the teacher to be minimal. The measures of the dependent variables were administered during the first two days (pretest) and last two days

(posttest) of the nine-week period. A research assistant administered both pretests and posttests.

Because the four classes involved were intact groups, the study employed a "nonequivalent control group design" as described by Campbell and Stanley (1966: 40). In this design, differential treatment effects are inferred from differential pretest-posttest gains for the various groups. In the repeated-measures ANOVA (the mode of analysis used in the present study), a treatment effect would be indicated by a significant treatment-by-time interaction with the experimental classes gaining more than the control classes. Similarly, differential effectiveness of the treatment for the two ability levels would be reflected in a treatment-by-ability-by-time interaction.

Another way to anlyze date from a nonequivalent control group design which complements the repeated-measures ANOVA has been suggested by Cronbach and Furby (1970). The technique involves calculating the regressions of the posttest on the pretest for each of the four classes involved. The ANOVA tells us how the classes as a whole did relative to each other: the regression lines indicate how students with various levels of pretest achievement did on the posttest within each treatment group. The latter information was judged especially useful since the grouping of classes in terms of low and average ability was based on gross measures of ability and thus not likely to be as sensitive to treatment-by-ability interactions as within-class regression lines.

EXPERIMENTAL TREATMENT

The experimental and control classes differed on two dimensions of classroom structure: the academic tasks which they performed and the mode of receiving feedback and reinforcement. The treatment differences on both dimensions are explained below.

Academic tasks: The students in the control classes listened to lectures, did math problem drills in class, and took three quizzes during the nine weeks of the experiment. The mathematics topics covered during the nine weeks were operations on fractions, decimals, and percents, in that order.

The experimental classes played the nonsimulation game EQUATIONS twice a week in addition to lectures, problem drills, and quizzes. The game was conducted in a fashion similar to that suggested by Allen (1969) and Allen et al. (1970). Within a classroom of thirty students, ten EQUATIONS games were played simultaneously (three students per game). The players at any given table were grouped homogeneously by math ability.

At the end of each game (a typical game lasted approximately five minutes), the scorer recorded the score of each player as specified in the rules. At the end of the class period, the game scores were summed to form a total-day score. Each day of play there was a winner at each one of the ten tables. Depending on whether a student won or lost at a table, he was moved up to a lower number table (tables were numbered 1 through 10) or down to a higher number table for the next day of play. This "bumping" procedure maintained homogeneous game tables while taking into account new learning (as reflected by a person winning or losing).

Reinforcement: The experimental and control classes also differed on the level of reinforcement contingencies and frequency of feedback. In the control classes, the students received feedback on daily drill exercises, which were corrected by class members at the end of a ten-minute drill session. In addition, they received grades on three hourly tests given during the nine-week period. All feedback received by the control students was contingent upon their individual performance.

In the experimental classes, the students received individual feedback on drills and quizzes similar to that received by the control classes. However, feedback on game performance was based upon group contingencies. The students were assigned by the teacher to four-member groups, with the specific intention of creating large within-group variation and small between-group variation. During the EQUATIONS game, each student acted as a "representative" of his team in competition with two other students, who represented other teams.

The group contingent feedback was administered via a newsletter that was distributed two days after each EQUATIONS tournament.[4] On the first page of the newsletter, the teams were ranked by their scores in the preceding tournament.[5] The teams were also ranked by their "season record," which was formed by adding the team's scores for all prior days of play. A commentary section of the newsletter congratulated the teams exhibiting good teamwork (assumed to be reflected in a general increase in the team members' performance). The second page of the newsletter listed, by team, the scores of the individual team members.

During the nine-week study, periodic practice periods were created in the experimental treatment during which teammates could work together. The teacher told the students that grades would be assigned on the basis of how well their team did in the game competition, but at no point were actual letter grades assigned.

DEPENDENT VARIABLES

The dependent variables consisted of three measures of mathematics achievement. The measures represented three levels of specificity of math skills that were likely to be affected by the experimental treatment. The first variable, representing the most general level of measurement, was the computations subtest of the Stanford Achievement Test in Mathematics. Measurement people have recently been critical of the use of standardized tests for curriculum evaluation (see Glaser and Nitko, 1971; Airasian and Madaus, 1972). Such tests, it has been noted, are designed for assessment of general achievement and thus are not sensitive to specific programs and short-term changes. But, the present study included such a test because standardized tests have the advantage of providing achievement norms which are readily understood by most educators. The computations subtest took one period (45 minutes) to administer. Using national norms, the raw scores were transformed to equivalent grade scores.

The second dependent variable represented a level of measurement specific to the subject matter covered by both treatments. It was obtained by counting only those items on the computations subtest which involved operations with fractions, decimals, or percents (the topics taught during the nine weeks of the study). A subject's score was the number right out of the 25 items so identified.

The third dependent variable, representing a level of measurement specific to the experimental treatment, was a divergent solutions test designed by the experimenters. The test consisted of two items. Each item specified a set of numbers and operations (called "resources") and the righthand side of an equation (called a "goal"). The students were asked to write as many different solutions as they could that would equal the goal, using only the numbers and operations given in the resources. A student's score was the total number of correct solutions written. The divergent solutions task is a basic skill used in playing EQUATIONS.

In terms of expected outcomes, the divergent solutions test would be most likely to reveal treatment differences, the content-relevant items from the computations subtest next likely and grade equivalents on the computations subtest least likely to reveal such differences.

RESULTS

The repeated-measures ANOVA on the divergent solutions test is given in Table 1. The between-subjects terms are included for completeness but are of little interest. For the within-subjects terms, the significant time effect ($F = 23.61$, $p < .001$) indicates that all students had increased in achievement over the nine-week period of the study; the students as a whole did learn. The significant treatment-by-time interaction ($F = 5.78$, p

$< .04$) reflects differential learning in the two treatment groups. Inspection of the means indicates that the experimental classes ($x_{Pre} = 6.82$, $x_{Post} = 9.87$) increased more than the control classes ($x_{Pre} = 6.25$, $x_{Post} = 7.32$). The two groups were approximately equal on the pretest. Figure 1 shows graphically the differential gains of the two groups on the divergent solutions test.

The significant treatment-by-ability-by-time interaction in Table 1 ($F = 4.53$, $p < .05$) indicates that the treatments resulted in differential learning for the two levels of ability. The graph in Figure 2 shows that the interaction was due to the large differential gains for the low-ability classes. The low-ability experimental class had the lowest pretest score but the second

TABLE 1
REPEATED MEASURES ANOVA ON THE DIVERGENT SOLUTIONS TEST

Source of Variance	df	MS	F
Between subjects treatment (A)	1	100.894 (η^2_B = .06)[1]	5.59*
Ability (B)	1	165.586 (η^2_B = .10)	9.17*
A x B	1	2.274	<1
$Error_B$	79	18.059	
Within subjects time (C)	1	166.000 (η^2_W = .21)	23.61**
A x C	1	40.654 (η^2_W = .05)	5.78*
B x C	1	2.009	<1
A x B x C	1	31.821 (η^2_W = .04)	4.53*
$Error_W$	79	7.032	

* $p < .05$.
** $p < .001$.
[1] η^2_B = amount of between subjects variance explained.
[1] η^2_W = amount of within subjects variance explained.

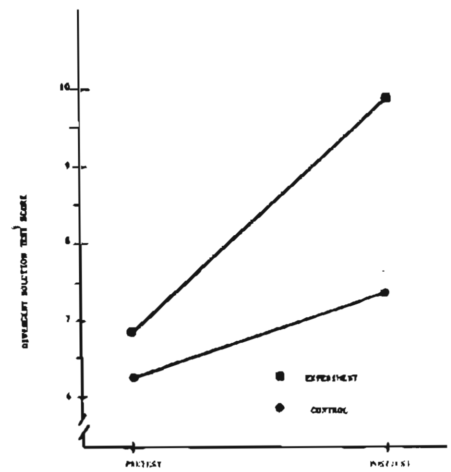

Figure 1: THE GAMES BY TIME INTERACTION ON THE DIVERGENT SOLUTIONS TEST

highest posttest score, while the low-ability control class showed virtually no gain.

The significant games-by-time and games-by-ability-by-time interactions accounted for a total of 9% of the within-subjects variation (5% and 4% respectively).

The results of the ANOVA for the content-relevant items of the computations subtest are given in Table 2. Again the within-subjects terms are the ones of interest. The significant treatment-by-time interaction ($F = 5.27$, $p < .03$) indicates differential changes for the two treatment groups over the nine weeks of the study. The means for the two groups show that the experimental classes ($x_{Pre} = 5.72$; $x_{Post} = 8.53$) had a larger

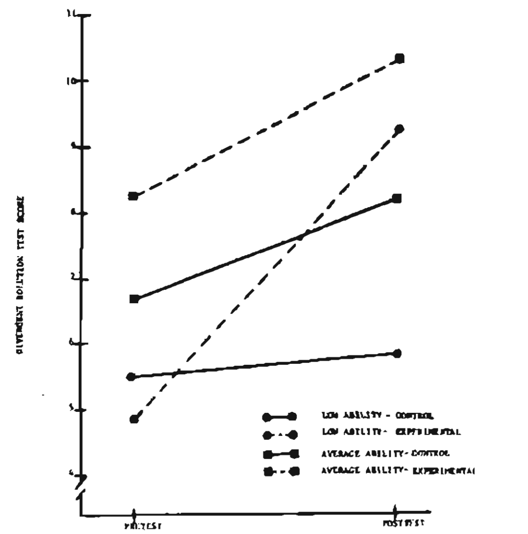

Figure 2: THE GAMES BY ABILITY BY TIME INTERACTION FOR THE DIVERGENT SOLUTIONS TEST

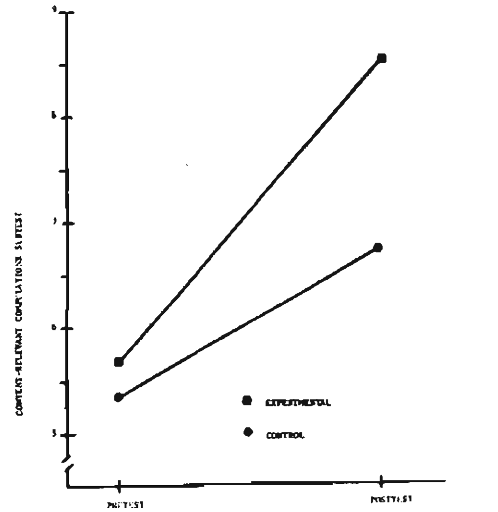

Figure 3: THE GAMES BY TIME INTERACTION FOR CONTENT-RELEVANT ITEMS ON THE COMPUTATIONS SUBTEST

gain than the control classes (x_{Pre} = 5.37; x_{Post} = 6.78). The significant treatment-by-time interaction is shown graphically in Figure 3. The treatment-by-ability-by-time interaction was not significant for the content-relevant items.

The analysis of variance for the computations subtest grade equivalents is given in Table 3. The results parallel those given in Table 2. The treatment-by-time interaction is significant (F = 4.66, $p < .04$) with the experimental classes (x_{Pre} = 5.5; x_{Post} = 6.3) showing a significantly larger gain than the control classes (x_{Pre} = 5.4; x_{Post} = 5.7). The significant effect is shown graphically in Figure 4.

TABLE 2
REPEATED MEASURES ANOVA ON CONTENT RELEVANT ITEMS FROM THE COMPUTATIONS SUBTEST

Source of Variance	df	MS	F
Between subjects treatment (A)	1	53.53 (η^2_B = .02)[1]	3.26
Ability (B)	1	449.88 (η^2_B = .22)	27.37**
A x B	1	1.64	<1
$Error_B$	92	16.43	
Within subjects time (C)	1	210.422 (η^2_W = .32)[2]	47.16**
A x C	1	23.522 (η^2_W = .04)	5.27*
B x C	1	6.896	1.54
A x B x C	1	9.183	2.06
$Error_W$	92	4.461	

* $p < .05$. [1] η^2_B = amount of between subjects variance explained.
** $p < .0001$. [1] η^2_W = amount of within subjects variance explained.

The pretest and posttest means, as well as the gains for each of the four classes on the three measures of achievement, are given in Tables 4, 5, and 6. The low-ability experimental class had the largest gains for each of the three dependent variables.

The lines showing the regression of the posttest on the pretest for each of the four classes on the solutions test and the computations subtest content-relevant items are given in Figures 5 and 6 respectively. The regressions for the computations subtest grade equivalents were very similar to Figure 6 and are omitted. In interpreting the regression results, two points need to be made. Treatment effects are reflected in the distance between lines. Treatment-by-ability interactions are reflected by non-parallelism of regression lines, as well as by the relative distance between the two experimental lines and the two control lines. The large black dots in the figures are points whose coordinates are the pre- and posttest means for the particular treatment group. The 95% confidence intervals for the regression slopes for the four classes on the three measures of achievement are given in Figure 7; the confidence intervals for the intercepts are given in Figure 8.

TABLE 3
REPEATED MEASURES ANOVA ON COMPUTATIONS SUBTEST

Source of Variance	df	MS	F
Between subjects treatment (A)	1	5.77	3.18
Ability (B)	1	56.88 (η^2_B = .25)[1]	31.36**
A x B	1	.62	<1
$Error_B$	92	1.81	
Within subjects time (C)	1	13.921 (η^2_W = .19)[2]	23.57**
A x C	1	2.751 (η^2_W = .04)	4.658*
B x C	1	.098	<1
A x B x C	1	.067	<1
$Error_W$	92	.591	

* $p < .05$. [1] η^2_B = amount of between subjects variance explained.
** $p < .0001$. [1] η^2_W = amount of within subjects variance explained.

DISCUSSION

Our general conclusion is that combining the nonsimulation game EQUATIONS with team competition significantly increased students' mathematics achievement over that of a traditionally taught class. The effect was observed for skills specific to the game as well as more general arithmetic skills.

The treatment-by-ability interaction for the divergent solutions test was due to the large gain by the low-ability experimental class; the mean on the posttest was almost double that of the pretest. Since this group had the lowest pretest score, one might argue that they would have the largest expected gain on the basis of regression effects alone. However, the size of the gain for this group, the lack of a gain for the low-ability control class, and the large gain for the average-ability experimental class (which had the highest pretest score) all indicate the results are not regression artifacts. Apparently average-ability students are able to do better on the creative mathematics task of making divergent solutions after learning more math, though playing the game results in a larger increase.

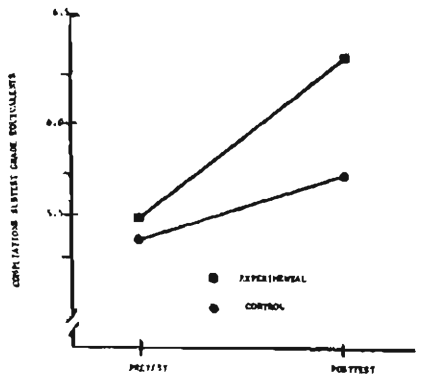

Figure 4: THE GAMES BY TIME INTERACTION FOR GRADE EQUIVALENTS ON THE COMPUTATIONS SUBTEST

TABLE 4
PRETEST, POSTTEST, AND GAIN SCORE MEANS FOR NUMBER CORRECT ON SOLUTIONS TEST

Ability Level	Treatment	Pre	Post	Gain
Low	Control	5.53	5.82	.29
	Experimental	4.88	9.24	4.36
Average	Control	6.70	8.26	1.56
	Experimental	8.32	10.36	2.04

TABLE 5
PRETEST, POSTTEST, AND GAIN SCORE MEANS FOR SCORE ON CONTENT-RELEVANT ITEMS FROM COMPUTATIONS SUBTESTS

Ability Level	Treatment	Pre	Post	Gain
Low	Control	4.19	4.67	.48
	Experimental	3.68	6.58	2.9
Average	Control	6.29	8.32	2.03
	Experimental	7.11	9.86	2.75

TABLE 6
PRETEST, POSTTEST, AND GAIN SCORE MEANS FOR GRADE EQUIVALENTS ON THE COMPUTATIONS TEST

Ability Level	Treatment	Pre	Post	Gain
Low	Control	4.85	5.17	.32
	Experimental	4.75	5.64	.89
Average	Control	5.85	6.15	.30
	Experimental	6.06	6.78	.72

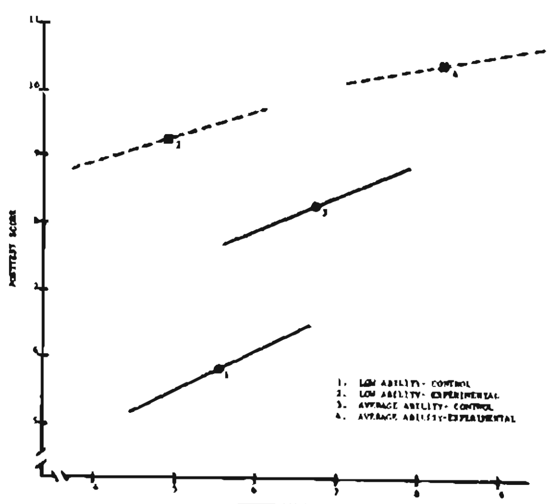

Figure 5: PRETEST-POSTTEST REGRESSION LINES WITHIN CLASSES FOR THE DIVERGENT SOLUTIONS TEST

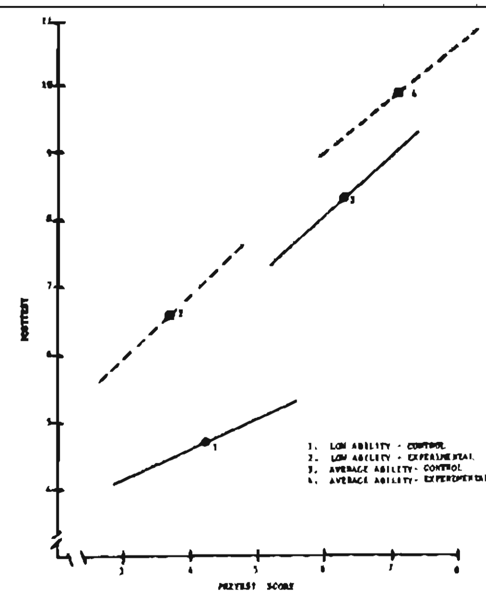

Figure 6: PRETEST-POSTTEST REGRESSION LINES WITHIN CLASSES FOR THE CONTENT-RELEVANT ITEMS OF THE COMPUTATIONS SUBTEST

whereas, the low-ability students require practice with the specific skill which the game provides in order to show any increase in the skill. There is a need to examine the effects of the game on other measures of math creativity to see if the skill generalizes to any degree.

The more traditional test of math achievement, the Stanford Achievement Test, also revealed a significant effect for the experimental treatment. The treatment by ability interactions for the computations subtest variables were not significant. However, for both the content-relevant items and the grade equivalents, the low-ability experimental class had the largest gains. The fact that the gains in grade equivalents was largely due to the gains on content-relevant items of the computations test is acknowledged by the authors. It was of interest to note that the significant game effect was detected even after the total raw scores (sum of relevant and nonrelevant items) had been converted to grade equivalents.

The striking fact about the regression analyses was how similar the regression lines were for the low- and average-ability experimental classes and how different the two regression lines were for the low- and average-ability control classes. The contrast was true for both the slopes and intercepts and was greatest for the grade equivalents variable. For example consider the regressions in Figure 6. The predicted posttest scores in each of the four groups for a student with a pretest

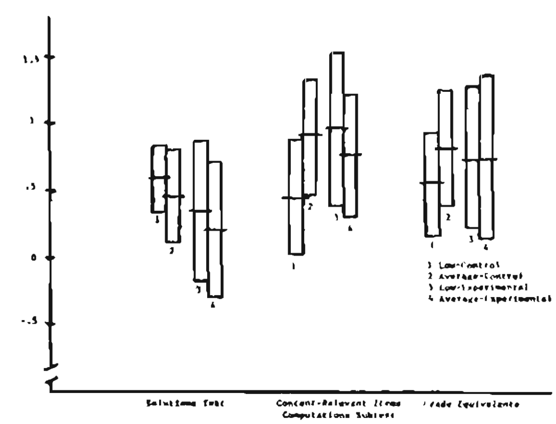

Figure 7: 95% CONFIDENCE INTERVALS FOR REGRESSION SLOPES

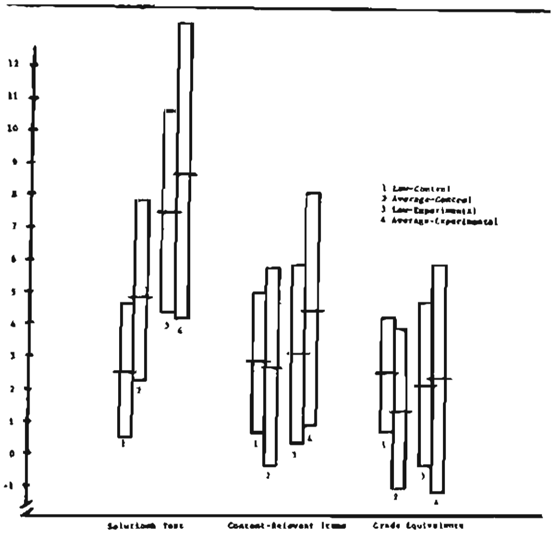

Figure 8: 95% CONFIDENCE INTERVALS FOR REGRESSION INTERCEPTS

score of 5.5 were 5.3 (low-control), 7.6 (average-control), 8.3 (low-experimental), 8.6 (average-experimental). The posttest score was more dependent on ability level of the class for the control classes than for the experimental classes. The games tended to reduce the differential learning rates evident in classes of different ability levels.

As indicated previously, the design of the study did not allow us to determine which aspects of the game classes resulted in significantly greater learning. The classroom teacher noted several key aspects of the experimental classes. First, the game succeeded in "turning on" students who had not been putting forth any effort. Second, the team competition feedback to the students via the newsletter was important to them. On two occasions when the experimenters failed to return the newsletters to the teacher on time, she reported that the classes became upset because they wanted to see how their teams had done. Third, during the periods of game play, the teacher found it much easier to give help to individual students while the rest of the class kept busy with the game. Fourth, students had a reason to help fellow team members to improve their teams' performances. Fifth, the game competition gave the students a reason to learn mathematics (to win the game). It is likely that the other classroom instruction benefited from this motivation. Sixth, playing the game was involving and fun.

It is likely that all these elements combined to produce the significant results obtained; thus, generalizations to games not structured like EQUATIONS (or even to EQUATIONS without team competition) are not warranted. The authors have completed a study which examines the independent and combined effects of games and teams on classroom process and attitudes as well as on achievement (Edwards and DeVries, forthcoming).

The present study and that of Allen et al. (1970) provide clear evidence of the effectiveness of EQUATIONS-type nonsimulation games.[6] The results of the present study provide support for a technique whose advocates have been hard-pressed to produce specific evaluative data.

1. EQUATIONS is a nonsimulation game in that no attempt is made to simulate some aspect of reality in either the game structure itself or the roles assigned to the players. Simulation games, on the other hand, do attempt to simulate some aspects of reality by either or both of these means.

2. The authors recognized that the challenge system could result in incorrect solutions being accepted as correct. A pilot test of the game revealed that such a situation arises infrequently and is most likely to occur among players with minimal mathematics skills. The teacher was alerted to this possibility, and consequently monitored closely such low-ability students.

3. Both probability and immediacy of reinforcement have been shown to be critical factors in inducing systematic behavioral change (compare Skinner, 1969). Each game of EQUATIONS can result in one or two winners, depending on the outcome of a challenge.

4. Allen et al. (1970) suggest the use of a newsletter as an effective medium for informing students of their performance, and at the same time generating excitement about the treatment.

5. The team score was obtained by summing the scores of the team members present for the game. If one or more team members were absent, the team was not compensated.

6. Stanley (1971) has raised a number of questions concerning possible sources of invalidity for the Allen et al. (1970) study because of the nonequivalent control group design employed. Since the same design was used in the present study, his points are dealt with here. First, the pretest showed that classes at the same ability level were not too different and thus differential susceptibility to change was unlikely. A treatment by testing interaction for any of the reasons mentioned by Stanley did not seem to apply to the present study. Both the pretests and posttests were administered by a research assistant under similar classroom conditions and time constraints. Finally, a novelty effect of the games for both the teacher and students seems inadequate to account for the significantly greater achievement gains observed in the two games classes over the nine weeks of the study.

REFERENCES

AIRASIAN, P. W. and G. F. MADAUS (1972) "Criterion-referenced testing in the classroom." NCME Measurement in Education 3: 1-8.

ALLEN, L. E. (1969) Equations–The Game of Creative Mathematics. New Haven: WFF N' PROOF.

--- R. W. ALLEN, and J. ROSS (1970) "The virtues of nonsimulation games." Simulation and Games 1: 319-326.

BOOCOCK, S. S. and E. O. SCHILD [eds.] (1968) Simulation Games in Learning. Beverly Hills: Sage Pubns.

BRASKAMP, L. A. and R. M. HODGETTS (1971) "The role of an objective evaluation model in simulation gaming." Simulation and Games 2: 197-212.

BRONFENBRENNER, U. (1970) Two Worlds of Childhood. New York: Russell Sage Foundation.

CAMPBELL, D. T. and J. C. STANLEY (1966) Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally.

CHERRYHOLMES, C. H. (1966) "Some current research on effectiveness of educational simulations: implications for alternative strategies." American Behavioral Scientist 19 (October).

COLEMAN, J. S. (1972) "Coleman on the Coleman Report." Educational Researcher 1: 13-14.

--- (1959) "Academic achievement and the structure of competition." Harvard Educational Rev. 29: 339-351.

CRONBACH, L. J. and L. FURBY (1970) "How to measure change–or should we?" Psych. Bull. 74: 68-80.

EDWARDS, K. J. (1971a) "Students' evaluations of a business simulation game as a learning experience." Johns Hopkins University Center for Social Organization of Schools Report 121.

--- (1971b) "The effect of ability, achievement, and number of plays on learning from a simulation game." Johns Hopkins University Center for Social Organization of Schools Report 115.

--- and D. L. DeVRIES (forthcoming) "The effectiveness of games and teams for teaching mathematics." Johns Hopkins Center for Social Organization of Schools Report.

FARRAN, D. C. (1968) "Competition and learning for under achievers." pp. 191-224 in S. S. Boocock and E. O. Schild (eds.) Simulation Games in Learning. Beverly Hills: Sage Pubns.

FLETCHER, J. L. (1971) "Evaluation of learning in two social studies simulation games." Simulation and Games 2. 259-286.

GLASER, R. and D. J. KLAUS (1966) "A reinforcement analysis of group performance." Psych. Monographs, General and Applied 80: 1-23.

GLASER, R. and A. J. NITKO (1971) "Measurement in learning and instruction." pp. 625-670 in R. L. Thorndike (ed.) Educational Measurement. Washington, D.C.: American Council on Education.

HAMBLIN, R. L., D. BUCKHOLDT, D. FERRITOR, M. KOZLOFF, and L. BLACKWELL (1971) The Humanization Processes. New York: Wiley-Interscience.

LEE, R. S. and A. O'LEARY (1971) "Attitude and personality effects of a three-day simulation." Simulation and Games 2: 309-348.

LIVINGSTON, S. A. and C. S. STOLL (1972) Simulation Games: An Introduction for the Social Studies Teacher. New York: Free Press.

SCHILD, E. O. (1968) "The shaping of strategies," pp. 143-154 in S. S. Boocock and E. O. Schild (eds.) Simulation Games in Learning. Beverly Hills: Sage Pubns.

SKINNER, B. F. (1969) Contingencies of Reinforcement: A Theoretical Analysis. New York: Appleton-Century-Crofts.

SLATER, P. E. (1965) "Contrasting correlates of group size," in I. D. Steiner and M. Fishbein (eds.) Current Studies in Social Psychology. New York: Holt, Rinehart & Winston.

SPILERMAN, S. (1971) "Raising academic motivation in lower class adolescents: a convergence of two research traditions." Sociology of Education 44: 103-118.

STANLEY, J. C. (1971) "Do non-simulation games raise I.Q.'s, whereas simulation games do not?" Simulation and Games 2: 119-122.

WODARSKI, J. S., R. L. HAMBLIN, D. BUCKHOLDT, and D. FERRITOR (1971) "Effects of individual and group contingencies on arithmetic performance." Presented at the meetings of the American Educational Research Association, New York.