DOCUMENT RESUME

ED 287 591 PS 016 912

AUTHOR Longstreth, Langdon E.; Alcorn, Mark B.
TITLE Practice: How Similar Must It Be To Affect Block Design Scores?
PUB DATE [87]
NOTE 22p.; The last page may not reproduce well.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Achievement Gains; *Games; *Intelligence Tests; Kindergarten Children; Performance Factors; *Preschool Children; Preschool Education; Primary Education; *Test Results; *Transfer of Training
IDENTIFIERS Critical Analysis

ABSTRACT

It was recently concluded that experiments "indicate that relatively brief interactions with a commercial game can cause a significant improvement in children's performance on an IQ subtest" (Dirks, 1982). The present paper makes three points. First, Dirks' analyses were not appropriate to the research question, and hence provide no rigorous evidence justifying the conclusion. Second, even if the analyses had been appropriate, the most the experiment could possibly have demonstrated was that a game nearly identical to the Wechsler Intelligence Scale for Children--Revised (WISC-R) Block Design subtest could elevate subjects' scores on the subtest. Third, by employing two games of less but still substantial similarity to block design tasks, the present experiment found no evidence of positive transfer to the Block Design Subtest of the Wechsler Preschool Primary Scale of Intelligence (WPPSI). Taken together, it would appear that, for the present, there is no rigorous evidence that games of any degree of similarity other than identity (test-retest effect) affect scores on the WISC-R or WPPSI subtests. (Author/RH)

Practice: how similar must it be to affect Block Design scores?

by

Langdon E. Longstreth
University of Southern California

Mark B. Alcorn
University of Northern Colorado

Running Head: Practice and Block Design scores

ABSTRACT

It was recently concluded that experiments "indicate that relatively brief interactions with a commercial game can cause a significant improvement in children's performance on an IQ subtest" (Dirks, 1982, p. 109). The present paper makes three points. First, Dirks' analyses were not appropriate to the question, and hence provide no rigorous evidence justifying the conclusion. Second, even if the analyses had been appropriate, the most the experiment could possibly have demonstrated was that a game nearly identical to the WISC-R Block Design subtest could elevate its score. Third, by employing two games of less but still substantial similarity, the present experiment finds no evidence of positive transfer on the Block Design subtest of the WPPSI. Taken together, it would appear that, for the present, there is no rigorous evidence that games of any degree of similarity other than identity (test-retest effect) affect scores on the WISC-R or WPPSI subtests.

The present experiment is concerned with the effect of play behavior on related IQ-test item performance. The impetus for the study comes from two contrary themes. On the one hand, experimental intervention studies during the preschool years have not been notably successful in changing IQ-test scores, not even the more prestigious large-sample longitudinal studies (e.g., Lazar and Darlington, 1982). On the other hand, there are reports that home play experiences do affect these scores (e.g., Stevenson, Parker, Wilkinson, Bonnevaux, and Gonzalez, 1978). Among the latter studies is one that reports what at first glance looks like a startling result, that 30 minutes of exposure to a commercially-produced game similar to the Block Design subtest of the WISC-R results in significant improvement on that subtest (Dirks, 1982). Closer reading, however, reveals that the game, "Trac 4", is almost identical to the Block Design subtest. That is, red and white cubes, with some sides divided along the diagonal, are to be sorted to match a geometric pattern depicted in a pictured square model. Thus the same material (blocks) with the same colors (red and white) with the same patterns (horizontal and diagonal color divisions) are used to seek the same goal (pattern matching). Furthermore, response time is measured in both tasks. So it is possible to conclude simply that 30 minutes of practice on a Block Design-like game improves performance on the Block Design subtest.

This is not a particularly surprising finding. An interesting extension of the study would be to vary the similarity of the game and the test item, and thereby to get some idea of the functional relationship between game-test similarity and test-score gain. The Block Design subtest is a particularly good choice for such a study, because there is considerable evidence that test items which require discovery of figural relationships are more highly g-loaded than any other type of item (e.g., Jensen, 1980, p. 229). In the present study, then, block play was provided to two experimental groups, and pre-post test scores were compared to those of a control group. The two experimental groups varied in the similarity of the block material to the test material.

METHOD

Subjects and IQ test. Subjects were 37 children with ages between 3 years, 10 months and 6 years, 2 months, with a mean of 59.4 months. All subjects attended a day-care center operated by the University of Southern California. Ethnic background was diverse, and the families in general were middle class. The 19 girls and 18 boys comprising the sample were selected with two criteria in mind: (1) they expressed a desire to "play games" with the experimenter, and (2) they met the age restrictions of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI).

With preschool children, the appropriate Wechsler IQ test is the Wechsler Preschool and Primary Scale of

Intelligence (WPPSI). The Block Design subtest of the WPPSI is similar to that of the WISC, with a few minor differences, the largest of which is the use of 2-dimensional blocks instead of 3-dimensional blocks. Other ingredients are approximately the same: red and white colors divided on the horizontal and diagonal, instructions to copy a design, speed instructions, etc.

Design. The subjects were randomly assigned to three groups, two experimental groups (N = 11 and 12), and a control group (N = 14). One experimental group was exposed to red and white blocks with at least one side of each block diagonally divided into red and white sections. These patterns are the same as the patterns on the block design subtest of the WPPSI. The other experimental group was exposed to similar blocks, but they were blue and yellow, and were divided horizontally instead of diagonally. The control group was not exposed to blocks at all, but rather to typical preschool material such as color books, etc. The purpose of this group was to control for other factors (familiarity with the experimenters, practice at following directions, etc.) which might possibly produce positive transfer.

Materials. The critical training stimuli were the blocks, which were made of wood (small blocks) or cardboard (large blocks). They were non-uniform in shape and size, their shapes being either square or rectangular, and their dimensions ranging from about 1 in per side to 12 in per side. Blocks for the red-white experimental group were covered with red and white contact paper with at least one surface per block divided diagonally into red and white sectors, and other surfaces being either red or white. Blocks for the blue-yellow experimental group were covered with blue and yellow contact paper with at least one side divided horizontally into blue and yellow sectors. There were 32 blocks for each condition. Materials for the control group were two commercial "workbooks" designed to prepare children for Kindergarten. These workbooks contained tracing shapes, geometric figures, coloring figures, and the like. So all groups were exposed to figural content, but only the two experimental groups manipulated blocks to match a standard.

Procedure. The two experimental groups were treated identically during play periods, except for the colors of their blocks. Subjects in each group were run in pairs in order to take advantage of peer-elicited activity and to increase the familiarity of the situation. Each subject participated in a one-half hour session, except that six of the 12 subjects from the red-white condition were given a second one-half hour play session on another day in order to explore training-duration effects. The Dirks study employed one one-half hour duration, split into 2 15-minute periods.

While no explicit instructions were given to the subjects, an experimenter was on hand to keep interest focused on the blocks and to urge constructive play; e. g., "Build me a castle" or "See how high a tower you can make". The experimenter occasionally participated in order to encourage more block play. The play area for the pairs of subjects was 8 x 10 ft, and no other toys were available during the play periods. The accompanying photograph shows one of the larger block constructions and its would-be architects.

Insert photograph about here

All testing and play sessions took place at the subjects' day-care center in rooms familiar to the subjects. There were three experimenters, all graduate students in child development. They had been extensively trained in administration of the WPPSI, and were familiar to the subjects as a result of free-play participation during the school year.

All subjects were pre- and postested on the five performance subtests of the WWPSI: (1) Animal House, a measure of learning ability to associate sign with symbol; (2) Geometric Design, a measure of ability to draw a series of pictures (i. e., circle, square, diamond); (3) Picture Completion, a measure of ability to identify missing parts of pictures of common objects; (4) Mazes, a measure of ability to solve pencil maze problems by tracing out an appropriate path; and (5) Block Design.

Pretests were administered approximately one month prior to exposure to the play sessions, and postests were administered 7 to 10 days later. The logistics of the situation made it impossible to schedule blind testing. However, experimenters did not necessarily postest the same subjects they had pretested, but rather about 1/3 of those subjects. Experimenters did not see the pretest protocals or scores after the initial scoring some 6 weeks earlier. They reported little recall, which they attributed to time and interference from intervening test administrations.

RESULTS

Table 1 presents means and standard deviations for scaled scores on all five performance subtests, separately for pretest and postest, and separately for treatment conditions. The two nested conditions within the experimental condition (one-half hour vs. one hour exposure) are collap d, as there was not the slightest suggestion of a difference in outcome.

Insert Table 1 about here

A glance at the grand means at the right column of Table 1 indicates a gain from pretest to postest for all five subtests. The cell contents show that this gain held for all experimental conditions, with no exceptions. A second effect indicated by the grand means is that the gains for Picture Completion and Blocks Design (about 2 units) are more than twice as large as the gains for any of the other three tests, for which the average gain is about 0.8 units. Third, scrutiny of the table contents reveals no other systematic patterns. In particular, it is not the case that the red-white condition produced the greatest gain for Block

Design.

These data were submitted to an unweighted-means analysis of variance (ANOVA) with experimental condition as a group factor (3 levels), and pre-postest (2) and subtests (5) as two repeated-measures factors. The results of this analysis are totally redundant with the observations just made: there were two significant effects, pre-postest gain, $F(1,34) = 48.3$, $p < .001$, and subtests by pre-postest interaction, $F(4,136) = 9.7$, $p < .02$. Of particular importance is the nonsignificant triple interaction, which reflects no treatment differences in relative gain among the five subtests, $F(8,136) < 1.0$.

These results can perhaps be more fully comprehended by looking at Figure 1, which presents the mean pre-postest difference for each treatment and for each subtest. It is clear from this figure that the red-white treatment did not produce a greater gain for Block Designs than for any other subtest, since there is a greater gain for Picture Completion. It is also clear that the red-white treatment did not produce the greatest gain for Block Designs, since the blue-yellow condition produced a greater gain. It is unlikely this result is due to chance group differences in ability to profit from the pretest administration, because the blue-yellow subjects gained less than the red-white subjects on two of the remaining four subtests.

Insert Figure 1 about here

Block Design scores by themselves were also submitted to ANOVA, with a single group factor (treatments) and a single repeated-measurement factor (pre-postest). If there were significant differences in treatment gain scores on the Blocks Design subtest, a significant interaction between the two factors would be expected. However, $F(2/34) = 1.4$; $p > .20$. So there is no indication that the three treatment conditions produced different gain scores for Block Design.

DISCUSSION

The present results suggest caution in assigning a causative role to game experience in boosting IQ-scores when the game-subtest similarity is less than near perfect. The present study found no evidence of transfer on the block design subtest when it was compared to transfer on other subtests unrelated to the experimental game.

Upon closer examination, it turns out that the Dirks study was also unable to produce relative positive transfer. Positive transfer was analyzed, but not positive transfer relative to that of other subtests. This omission is puzzling because the experimental design provided information for such an analysis. That is, pre-post test changes on the block design subtest as a function of game or no game experience should have been compared to the same changes on the Object Assembly subtest, the other subtest administered by Dirks. Such an ANOVA would have evaluated the three factors of game vs. no game, pretest vs. post test, and Block

Design vs. Object Assembly subtests. But two 2 x 2 ANOVAS are reported instead, one for each task.

The seriousness of this omission can be appreciated by a re-examination of Dirks' Figure 2, which presents the pre-test and post-test mean scores for the Block Design subtest alone (see our Figure 2, solid lines). A significant 2-way interaction is reported for these data, indicating that the post-test gain is greater for the game condition than for the no game condition. We have added to that figure the mean scores for the Object Assembly task (dashed lines). These means reveal the same 2-way interaction, although of lesser magnitude. What is not reported is the significance of the difference between these two interactions. Therefore we have no way of determining if the game activity improved Block Design scores more than Object Assembly scores or not. Such a conclusion would be warranted only if the triple interaction of all three factors had been found to be significant. It was the absence of such an interaction in the present study that led us to accept the null hypothesis.

INSERT FIG 2 ABOUT HERE

We suggest, then, that there is little rigorous evidence that even a nearly identical game produces transfer specific to the Block Design subtest. Our own results replicate this non-finding, and extend it to games of lesser similarity. It may be the case, of course, that other games related to other subtests might produce positive transfer. It would seem, however, that at least for the WISC-R and the WPPSI, such a possibility remains to be demonstrated.

REFERENCES

Jensen, A. R. (1980). Bias in mental testing. New York: The Free Press.

Dirks, J. (1982). The effect of a commercial game on children's block design scores on the WISC-R IQ test. Intelligence, 6, 109-124.

Lazar, I. & Darlington, R. (1982). Lasting effects of early education: a report from the consortium for longitudinal studies. Monographs of the Society for Research in Child Development, No.. 195, 47, 1-151.

Stevenson, H. W., Parker, T., Wilkinson, A., Bonnevaux, B., & Gonzaiez, M. (1978). Schooling, environment. and cognitive development: a cross-cultural study. Monographs of the Society for Research in Child Development, No. 175, 43, 1-92.

Author identification footnote

We thank Margaret Howell and Colette Horn for their assistance in running subjects. Correspondence concerning this article should be addressed to Langdon E. Longstreth, Department of Psychology, Seeley G. Mudd Building, University of Southern California, Los Angeles, CA 90089-1061.

Table 1

Pre-postest scores on Five WPPSI Performance Scales for three Treatment Conditions

Test		Red-White		Blue-Yellow		Control		Grand
		Mean	SD	Mean	SD	Mean	SD	Mean
AH	Pre	12.4	2.3	12.5	1.8	12.4	2.6	12.4
	Post	13.5	3.5	13.3	1.5	13.7	3.1	13.5
PC	Pre	11.8	2.3	13.2	2.4	11.3	2.9	12.1
	Post	14.6	2.2	14.6	2.4	13.6	3.2	14.3
BD	Pre	12.9	2.4	12.1	2.8	13.5	1.7	12.8
	Post	14.9	1.8	14.7	2.9	14.6	2.3	14.7
M	Pre	12.8	1.9	12.7	1.7	12.6	1.7	12.7
	Post	13.3	1.7	14.1	2.5	12.9	1.4	13.4
GD	Pre	13.0	3.0	12.6	2.2	12.7	2.8	12.8
	Post	13.1	2.8	13.5	2.7	13.4	2.8	13.3

1. AH:Animal House; PC:Picture Completion; BD:Block Design; M:Mazes; GD:Geometric Design.

Figure captions

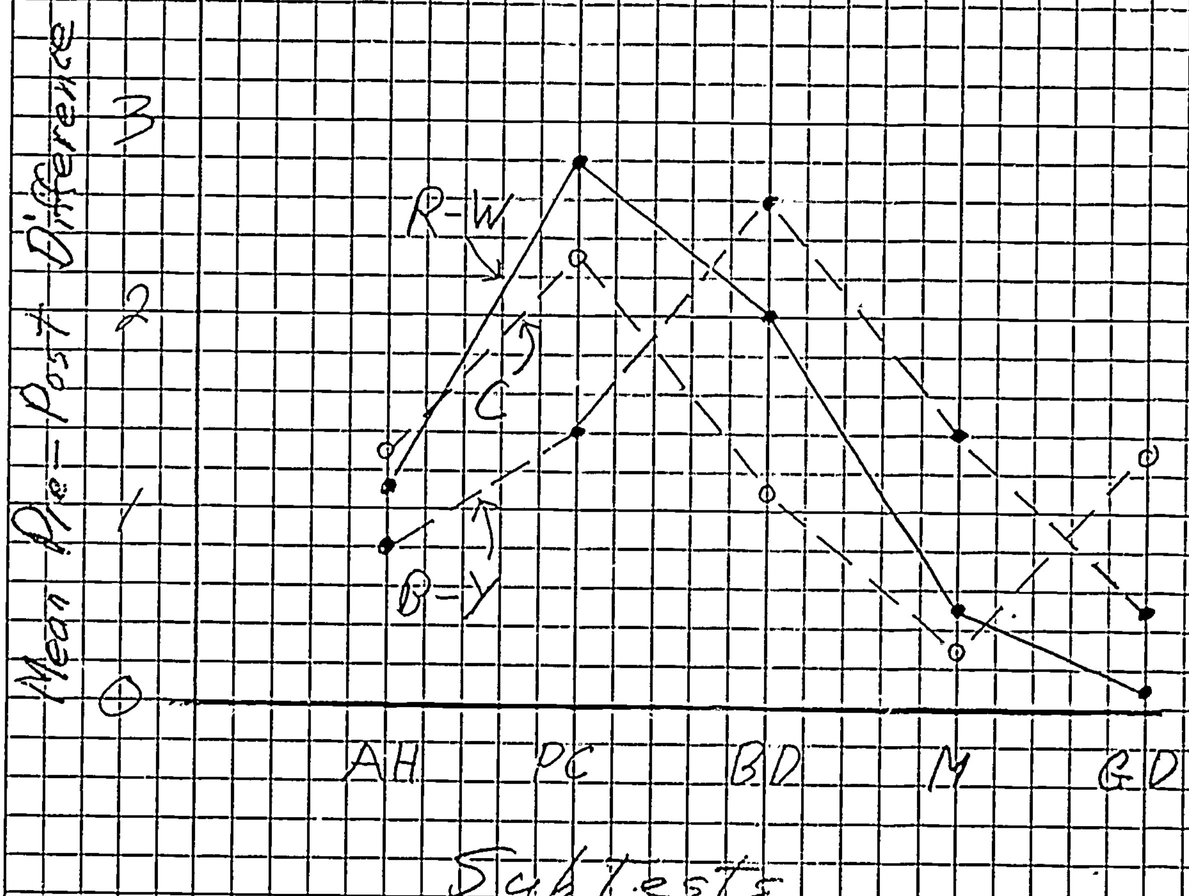
Figure 1. Mean pre-postest difference scores for two experimental groups and control group. AH: Animal House; PC: Picture Completion; BD: Block Design; M: Mazes; GD: Geometric Design.

Figure 2. Dirks' Figure 2, with addition of Object Assembly scores.

PHOTOGRAPH TITLE

Block play in the red-white experimental condition

Fig 1



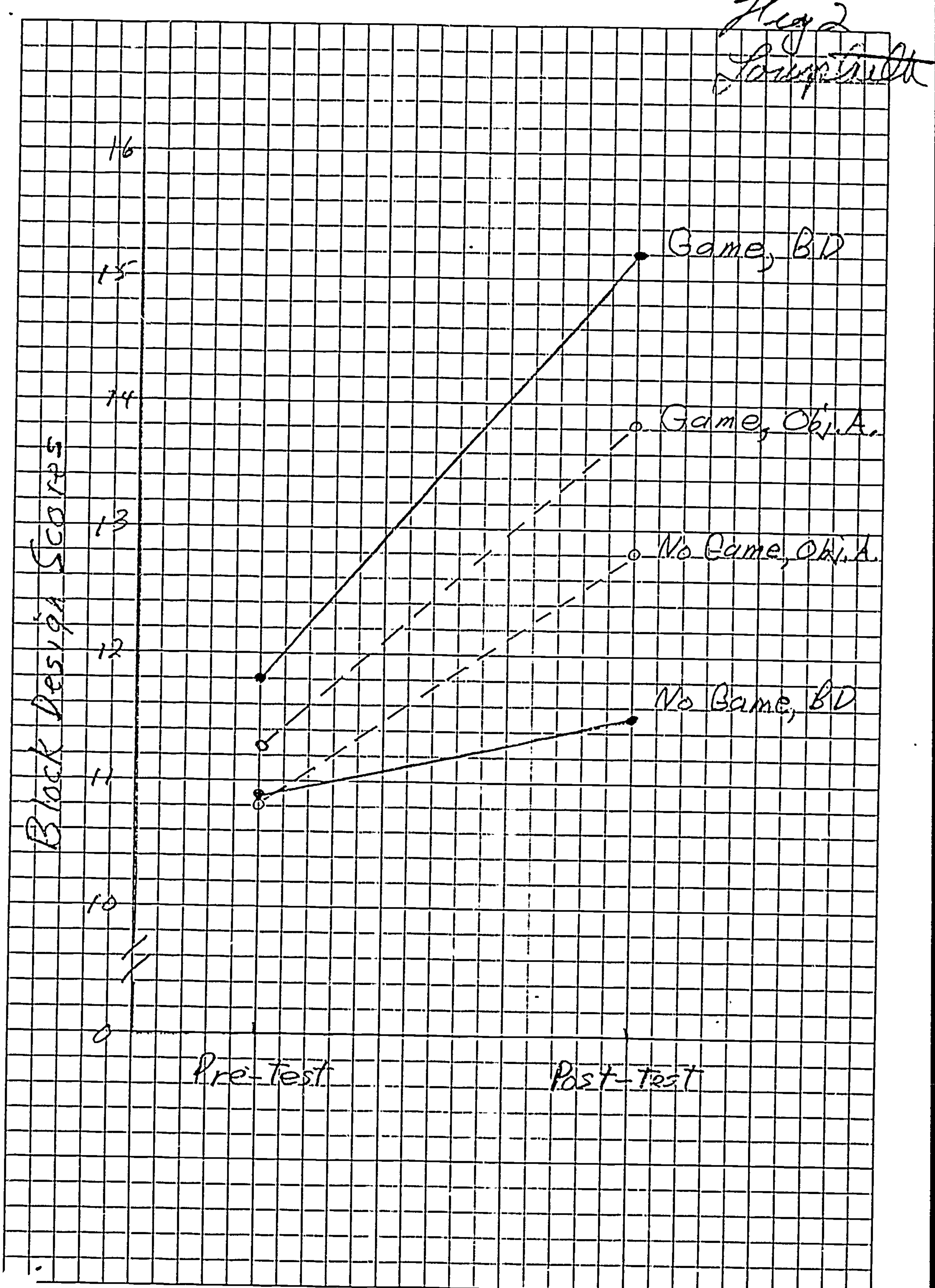


Fig 2
Block Design Scores
16
15
14
13
12
11
10
0
Game, BD
Game, Obj. A.
No Game, Obj. A.
No Game, BD
Pre-Test
Post-Test