

**VIRTUES OF A NON-SIMULATION GAME:
A Reply to Schild and Stanley's
Letters to the Editors**

The preceding comments of Julian Stanley and Erling Schild about the increase in nonlanguage I.Q. score accompanying intensive exposure to WFF 'N PROOF help to put into perspective the study that we reported in the September 1970 issue of *Simulation & Games*. We appreciate their interest and their suggestions for improving future investigation of the effects of instructional gaming.

Some background information may help to clarify our reasons for publishing the study at this time. When we became aware of the disappointing conclusions of the Cherryholmes survey, we felt that the results of some of our efforts (which had not been published and were not included in that survey) should probably be made more widely known to provide a more comprehensive picture of the full range of effects of instructional gaming. Our data were derived from the most careful evaluation that we could arrange with the available resources of an ongoing instructional gaming program in a regular summer school curriculum. The performances of the experimental and control groups were so different in some respects—and so alike in others—that it did not seem unreasonable to bring to the attention of those who are interested in this field what appeared to us as fairly strong evidence that something significant was occurring as a result of instructional gaming. The reported study represented the third time that an exposure to WFF 'N PROOF was accompanied both by an enormous increase in nonlanguage I.Q. scores and by a statistically insignificant change in language I.Q. scores. We believe that a better understanding

EDITOR'S NOTE: *The full data from the Allen et al. study will be published in the June 1971 issue of Simulation and Games, should anyone wish to redraw from it.*

of why these differences occur is important and may lead to a deeper understanding of the process of learning.

To supplement the evidence of no-change summarized in the Cherryholmes article, we decided to prepare a follow-up report for publication in the *American Behavioral Scientist*, where the Cherryholmes article had appeared, to indicate what evidence we had that some kinds of instructional gaming are accompanied by statistically significant differences in changes of scores on the California Test of Mental Maturity. Michael Inbar, editor of *Simulation & Games*, was then in the midst of arrangements for creating a new publication in the gaming field; when he learned that we were preparing such a report, he requested that we hold it for inclusion in the new publication to be launched in 1970 because it would be so clearly relevant for *Simulation & Games*.

A manuscript was prepared and submitted. In the course of interaction with the editorial staff, it was considerably shortened, and many of the details of the study, including the raw data, were deleted. We think that some of the questions raised by Stanley and Schild bear out our judgment that inclusion of the more complete summary of the data (which will be presented in the next issue) would have been preferable to the abbreviated summaries that ultimately appeared in the published report.

For example, it would have been clear that the differences hypothesized by Stanley between the experimental and control subjects as possible alternative explanations of the differences in nonlanguage score changes do not account for the results that occurred. He suggested that the advanced-level CTMM administered to the high school subjects might have been more subject to practice effect than the lower-level test is, and since proportionately more students in the experimental group were in high school than were those in the control group and thus took the advanced test, he further suggested that the experimental group may have benefited

more from the hypothesized difference in practice effect between the advanced- and lower-level tests. If the hypotheses were true, then the increases in scores for the high school students should be larger than the increases in scores for the junior high school students. However, the more complete summary of the data shows that for both the experimental and control groups the increases in nonlanguage I.Q. scores were larger for the junior high school subjects than for the high school subjects. The nonlanguage I.Q. score increase for the experimental junior high school students was 21.9 points whereas that for the experimental high school students was 16.6 points. Similarly, for the control junior high school students the increase was 7.0 points whereas for the control high school students it was only 2.0 points.

Another possible alternative hypothesis dispelled by the more complete data is that the experimental group was "brighter" and therefore better motivated to improve its scores on the retest. If brightness (as measured by the CTMM) were so associated with motivation, we should expect those in the experimental group with larger increases in non-language I.Q. score to have higher overall I.Q. scores. Again the data do not support the hypothesis. A breakdown of the data by quartiles shows no such relationship (see Table 1).

TABLE 1

Change Quartile	Average Increase in Scores – Non-Language	Average I.Q. Score
First	37.9	110.0
Second	23.7	114.3
Third	17.2	111.2
Fourth	4.4	113.5

On the possible differences between the experimental and control groups with respect to a boredom factor on the retest, we did not collect any data with respect to previous exposure to the test. We know of no reason why there should have been any difference in the groups in this respect, and

lacking evidence one way or the other we would assume no difference in this respect until some showing otherwise.

Concerning possible experimenter bias in administration of the CTMM to the experimental and control groups, the Vice Principal of the school, Nort Nelson, separately administered both the pre- and the post-tests to each group. Mr. Nelson taught the pre-algebra control group during the experimental three-week period. Two different teachers taught the experimental group, and they taught separately. Neither of the teachers of the experimental subjects participated in the testing in any way.

We appreciate the suggestions for alternative methods for conducting more systematic and rigorously controlled experiments when resources are available to so study the effects of instructional gaming. We concur wholeheartedly with the suggested usefulness of such studies to better understand what is happening in learning situations organized around such gaming. We particularly welcome suggestions about "appropriate indices of learning by conventional methods" and other indicators that are regarded as signaling significant change. In light of Stanley's suggestion particularly, it is perhaps worthy of mention in passing that in a similar earlier study the arithmetical reasoning component of the non-language part of the CTMM was the one in which the largest increases in scores occurred.

We recognize the rather casual nature of the reported study, but felt that its results might offer useful encouragement to many persons interested in educational innovation and would present a more rounded picture to any who might have been misled by the negative conclusions of the Cherryholmes review. Our report was certainly not intended to sound like the last word on this subject; rather, it was intended more as a stimulus to further research.

LAYMAN E. ALLEN
University of Michigan

ROBERT W. ALLEN
Nova University

JOAN ROSS
University of Michigan