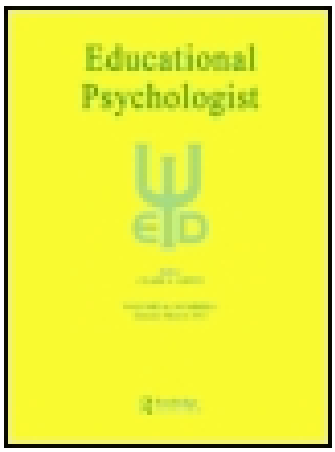


Educational Psychologist
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hedp20

# How we all failed in performance contracting

Ellis B. Page [a]
[a] University of Connecticut ,
Published online: 01 Oct 2009.

To cite this article: Ellis B. Page (1972) How we all failed in performance contracting , Educational Psychologist, 9:3, 40-42, DOI: 10.1080/00461527209529079

To link to this article: http://dx.doi.org/10.1080/00461527209529079

ments and evaluation policies. Third, students are encouraged to contact the Coordinator with questions on matters of detail, as registration, course content, module assignment, and room location. The Coordinator is also available to mediate student grievances should the occasion arise. Such a set of procedures has proven valuable in providing the student with a sense of security in the flexible modular program.

REFERENCES

Grinder, R. E. Flexibility and sequence: Educational psychology and the training of teachers. *Educational Psychologist,* 1970, *8,* 9–12.

Newsom, R. S. & Gaite, A. J. H. An evaluation of two approaches to the teaching of educational psychology. *Educational Psychologist,* 1971, *9,* 14–16.

Sattler, H. E., Grinder, R. E., & Clark, M. C. Prospects for modules in educational psychology. *Journal of Teacher Education,* 1971, *22*(4), 399–404.

## HOW WE *ALL* FAILED IN PERFORMANCE CONTRACTING[1]

Ellis B. Page
*University of Connecticut*

Educational psychologists little realize, so far, how our world has changed. Some of our most valued practices and assumptions have been put to a massive test, almost without our awareness, and have fared badly indeed. When we really grasp what happened, we will possibly never again feel quite so secure in our profession.

What happened was the extraordinary experiment conducted by the Office of Economic Opportunity (OEO). The subject was, at first glance, just one of many educational gimmicks currently on the scene: "performance contracting" (PC). Yet also on the block, because of the way the experiment was designed, were some common practices and beliefs about: (1) compensatory education; (2) the design of educational experiments; (3) the selection of suitable criterion tests; (4) the contributions of behaviorism to the schools. These are near the heart of our professional concerns, and those involved are our colleagues and brothers. And brothers, we have problems.

* * * * *

The OEO study of performance contracting is, quite possibly, the most impressive *experiment* ever conducted in education. As such, it has no competition from such massive non-experimental researches as Project Talent, National Assessment, and other status studies. In fact, it has very little competition at all, when one realizes how it was put together, and what a range of important subjects and situations it sampled.

The earlier trials of performance contracting, though flawed by poor controls and even by scandal, enjoyed such remarkable "success" that many believed that PC, in some form, might well be *the* solution to educational inequality. Among the believers were officials in OEO, who had smarted with the criticism that in Texarkana there was "teaching to the test," and who resolved to make, at last, a demonstration of PC which would be beyond reproach. Their faith was impressive:

> Staff visited Texarkana, and saw great promise in performance contracting as a means to *help poor children achieve the same results from classroom effort now achieved by nonpoor students.* (OEO, 1971, p. 3, italics added)

There in a phrase is the soul of compensatory education: poverty did the harm, and schooling will undo it all, and make it right again.[2]

It was not just the OEO who believed this; it was a vast number of contractors as well. If the contractors are now accused of overselling their product, it is unfair; they obviously believed their own claims. There were at least 50 companies bidding on work in 1971. Thirty-one of them bid on this major PC experiment, and six unlucky companies won the bid. It was a punishing experience for all six, and three or four are now out of the business. If there was a major deception, then, it was deep in the heart of our profession.

Listen to the OEO's description of the "relatively simple" system of PC:

> —A contractor signs an agreement to improve students' performance in certain basic skills by set amounts.
>
> —The contractor is paid according to his success in bringing students' performance up to those prespecified levels. If he succeeds, he makes a profit. If he fails, he doesn't get paid.
>
> —Within guidelines established by the school board, the contractor is free to use whatever instructional techniques, incentive systems, and audio-visual aids he feels can be most effective. He thus is allowed more flexibility than is usually offered a building principal or a classroom teacher. (OEO, 1971, p. 2)

Does it sound familiar? It should. It is at least implied by our behaviorist doctrine of the past several decades. It is main-line, established reinforcement theory put into action. Programmed instruction was strongly related to it, and behavior mod seems to have the action right now. One of its main assumptions is the crucial nature of *motivation.* And this problem may be overcome by incentives: e.g., "trading stamps, that they would use for gifts, and free time during the class period that they could use to read magazines, listen to records, or other recreational activities" (p. 3). Some of the anticipated gains would be "better overall performance ... accountability ... drop-out prevention ... integration ... individualization of instruction" (pp. 5–6).

To be important, an experiment must seek generality; and this remarkable experiment had it built into the design. The six PC contractors reflected the most persuasive and varied practices in applied psychology. Each hired company would apply its principles in three different school systems (18 in all), widely separated throughout the U.S., across six grade levels (1–3, 7–9), and across every major low-achieving ethnic group, from poor urban Anglos to Alaskan Eskimos. There was to be, furthermore, generalization across the two most important educational skills: reading and arithmetic. Over 25,000 students would participate, and over six million dollars would be spent.

Many intervention studies report glowing results to the newspapers, but upon examination are lacking in even the most primitive controls. But the OEO had been burned, and determined that this study was to be, within its realm of political possibility, a model of experimental practice. It came very close. The 18 participating districts were carefully chosen, and the subjects, both E and C, were from the worst-performing schools in the districts. And from each school, for each grade, the 100 students furthest behind their grade levels would serve as Ss. The resulting composite of subjects is very representative of the target population.

Precisely at this point of assignment to the E and C groups, where most wide-scale studies founder, the OEO likewise relaxed from the best research practice. Assignment of schools to treatments was very complicated in so widespread a study, where at least 100 E students were sought in each grade at each site, and where some schools obviously would not supply enough potential Ss. Therefore a number of compromises were made. But the most serious lapse from rigor, at least in theory, was permitting non-random considerations of cooperativeness, political considerations, and apparent need to decide which school was E and which C.[3] This resulted in a slightly abler group, on the average, for C schools as judged by the pretest.

As irritating as this compromise is to the professional psychometrician,[4] it is surely not fatal to this study. Battelle brought sophisticated attention to creating regression models which, under reasonable assumptions, adequately accounted for these group differences in the tests of main effects. This is a study which, because of its results, many would like to belittle or set aside. But with the widely heterogeneous Ss making up both E and C groups, with the broad generality of tests employed for contractual and evaluation measures, and the variety of experimental techniques applied, there is no justifiable claim that the E methods had substantial effects which were not evident in the analysis. The breadth of design weakens competitive hypotheses.

Even more than the other generality in the design, the dis-

tinctive achievement was in the criterion testing. Texarkana had left scars, and there would be, in this large PC study, no "teaching to the test" if OEO could help it.[5] These results were to be unassailable. The procedures were not simply those of ordinary good practice; they were virtually unprecedented.

To construct special tests would have been expensive, complicated, and questionable, and the decision was made to use most of the major, currently published tests in reading and arithmetic. These were to be given in an elaborate double-blind, with the identity of pre-tests and post-tests kept secret from everyone: contractors, teachers, and even students.

Furthermore, any attempt by the contractors to discover the identity of the test would be sufficient cause for the termination of the contract. And even the instructional content was to be spot-checked to make sure that no testing material was used in the conduct of instruction. Finally, tests were to be administered by outsiders, so that there would be neither identification of tests, nor aid to the students in taking them. And the evaluation of all the data was to be done by a responsible and disinterested corporation. In short, there was rare rigor and generality in the criterion testing, perhaps without parallel in large-scale research.

One remarkable feature of the design was the elimination of the usual upward drift caused by regression. The students were chosen *before* the pre-test on the basis of *other* data. Whatever general upward shift would occur from selection would already have happened *for the pre-test itself*. No upward change in the post-test general mean, then, would come as an unearned gift to the contractors. And the OEO had thus avoided one more of the most common artifacts in educational research.

* * * * *

Even with all this control, the OEO, the school systems, and the PC contractors went enthusiastically into the following arrangement: *Essentially, and in most cases, payment was to be made only for a student who gained one school year of growth between the fall pre-test and the spring post-test.*[6] Note, the design eliminated ordinary upward regression, teaching of the test, and biasing of the evaluation — all three of them the bane of sound inference, but the comfort of those determined to report favorable results. What was left was essentially improvement in the actual, general, *transferable* skills of reading and arithmetic.

How much improvement should one ordinarily expect in a year of schooling? From the lowest-achieving youngsters, given the usual instruction, one would expect perhaps .7 of a year — surely *not* a full year's growth. For as we know, in school ability the weak usually become weaker. That is, they maintain their relative position in standard deviations below the mean, and this means a greater and greater lag behind their abler classmates.

Given these warning signs, *why* did contractors, schools, and the OEO sign the contracts, and move so happily on toward their fates? The answer is inescapable: they *really believed* that the lagging pupils were disadvantaged only in their prior experience; that the public schools were terribly ineffective teachers; and that the application of the usual psychological principles would cause extraordinary leaps in achievement.

* * * * *

When the results were finally released, the shock was severe. The pupils fell far short of the year's gain required by contract, and the contractors were immediately in deep financial danger. And the OEO was in deep embarrassement. It had achieved a rare kind of research rigor, yet the testing results were shattering.[7] It was apparent that the hopes for PC which led to the contracts had been wildly optimistic.

There were educational scientists, of course, who believed that many gains claimed by compensatory education during the past decade resulted, in fact, from just such artifacts as the OEO was, at last, controlling: regression and practice effects, teaching to the test, biased analyses of data, and confusion of item samples with item domain. For such skeptics, it was no surprise that the contractors suffered loss, or that the target pupils failed to achieve a year in a year.[8]

Nevertheless, there is a deep shock for almost every educational psychologist who looks seriously into the data. The outcome is not simply that the PC enthusiasts were optimistic, but that *there was no detectable effect of the experiment.* In the words of the OEO (1972, pp. 17–24):

> [T]he difference in gains was remarkably small in all 10 of the grade/subject combinations for which this analysis is appropriate.[9] In half of the 10 cases, there was no difference at all between the gains of the experimental and control groups. In four of the cases there was a difference of as much as two-tenths of a grade level. These overall differences are so slight that we can conclude that performance contracting was no more effective in either reading or math than the traditional classroom methods of instruction.... [T]he performance of students in the experimental group does not appear disappointing just because students in the control group did unexpectedly well. In fact, neither group did well. In only two of the 20 possible cases was the mean gain of either the control or experimental students as much as one grade level.... In all cases, the average achievement level of children in the experimental group was well below the norm for their grade and in all cases, in terms of grade equivalents, the average slipped even further behind during the year.... [There] is no evidence that performance contracting had differential results for the lowest or highest achieving students in the sample.... Not only did both groups do equally poorly in terms of overall averages, but also these averages were very nearly the same in each grade, in each subject, for the best and worst students in the sample, and, with few exceptions, in each site. Indeed, the most interesting aspect of these conclusions is their very consistency. This evidence does not indicate that performance contracting will bring about any great improvement in the educational status of disadvantaged children.

No matter how these data are revolved (and more elaborate analyses are forthcoming), the essential conclusion will remain, and it seems a severe blow to certain of our professional illusions. Many of us have believed, implicitly, something like this:

> Applied psychology has certain powerful behavioral skills. We understand task analysis; input repertory; stimulus shaping; response elicitation; the provision of reinforcement; the arrangement of repetition, sequencing, looping; concept formation; the practicing of transfer. These are important ingredients in learning, and as psychologists we understand these things much better than traditionally trained teachers. If we as a profession are given the support, the students, the autonomy, we can make incalculable improvements in education.

This belief has been one cornerstone of our faith in ourselves.

Now the OEO has provided what may be the first really solid test of its truth: — whether the present, state-of-the-art, garden-variety, applied psychology can in fact contribute to the most important learnings in the schools. We will not make those statements so casually in the future. Our skills in training do *not* seem the immediate solution to our problems in education.

There are arguments to be raised, of course, some with validity: These PC contractors had materials and programs that were far from ideal. This is only a limited test of one year's duration. This study compared a well-bred horse (teaching) with the *first* automobile (PC). The tests are perhaps biased toward the conventional methods of teaching. We have not seen the affective changes. Perhaps the higher-order interactions...

Yet this experience with performance contracting, painful as it is, should be squarely confronted. It may cause us to re-examine many of our practices and assumptions about school learning and its evaluation. We have much to reappraise. It will be exciting to see how, in the next few years, we cope with this threat to our professional confidence.

## REFERENCES

Battelle Memorial Institute. The Office of Economic Opportunity experiment in educational performance contracting. Research report, January 29, 1972. Columbus Laboratories.

Campbell, D. T. and J. C. Stanley. Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of Research on Teaching.* Chicago: Rand McNally, 1963. Pp. 171–246.

Office of Economic Opportunity. Performance contracting: An experiment in accountability. OEO Pamphlet, 1971? Washington, D.C.

Office of Economic Opportunity. An Experiment in performance contracting: Summary of preliminary results. OEO Pamphlet 3400–5, February, 1972. Washington, D.C.


## FOOTNOTES

1. Thanks are given to Charles Stalford of the OEO, to Jason Millman, and to Bruce Rogers for their own opinions and comments. The present writer did a short consulting assignment with OEO concerning this project, but had no part in the design or analysis of the study. And this article is based on sources which are publicly available. Naturally, no one else should be considered party to the opinions or conclusions expressed here.

2. The official current position of the OEO, notwithstanding such quotes as given, is that it was not *promoting* PC as a solution, but was *testing* it as one promising effort. And the rigor of the test surely lends weight to this position.

3. According to Battelle (1972, p. 8): "Generally speaking, the most deficient school or schools were selected as the experimental school(s) and the next most deficient as the control schools. In large districts where there was sometimes a substantial choice of schools, the selection was affected by the presence or absence of other special programs in them, the receptivity of individual schools to being included in the experiment and the relative efficiency with which the required number of experimental and control subjects could be accumulated in them."

4. As Campbell and Stanley (1963) noted, randomization is the *only* all-purpose method of achieving pre-treatment equality.

5. Such test-teaching is difficult to control. In another study, one PC contractor held the position that the target *domain* of content was precisely the *six reading tests* which might be employed as criterion tests! Thus he developed a drill device to teach the vocabulary *from those tests,* and maintained that this was acceptable practice, despite the objections of testing experts.

6. Since the contractual arrangements were based on the E group's own progress, regardless of the controls, the biased assignment to E and C would not importantly affect payment.

7. It is possibly unrelated, but there has been considerable shifting in key OEO personnel concerned with the PC experiment.

8. Recently, a study very popular with the press has raised the IQs of disadvantaged pupils by "an average of 30 points." Repeated requests have so far failed to elicit a technical report which would permit detailed examination.

9. "Gain" scores were difficult to calculate or rationalize for the first grade work.

# DIVISION 15
# EXECUTIVE COMMITTEE MEETING

April 5, 1972

*Palmer House*
*Chicago*

Present: Joe L. Byers, Arthur P. Coladarci, John Feldhusen, Robert Glaser, Robert E. Grinder, Robert J. Havighurst, Herbert J. Klausmeier, Frederick J. McDonald, S. Jay Samuels, Joanna Williams

The meeting was convened at 4:00 p.m. by President Glaser.

### *COMMITTEE REPORTS*

*Nomination Committee.* Arthur Coladarci, Chairman of the Nomination Committee, submitted the Committee's slate of nominees to the Executive Committee for the offices of Division President, Secretary-Treasurer and member-at-large of the Executive Committee.

Dr. Coladarci raised a question from his Committee regarding the by-laws of the Division for nominations. As stated in the by-laws, "each year a call for suggested nominees for offices, the terms of which expire, shall be issued to members of the Division. The determination of the basis for selecting the final slate of nominees is as follows: the two people who are most frequently nominated by the membership are placed on the ballot; in addition, the Nomination Committee may select one or two more nominees from among the remaining nominations by members or from among members of Division 15 who are eligible and willing to serve." In operation, Dr. Coladarci pointed out, because of a low rate of response from the members on the call for nominations, the two who are most frequently nominated by the members are placed on the ballot with sometimes as few as only five people nominating them from the total division membership. The Executive Committee decided that the by-laws regarding nominations should remain in operation for another year to see if this continues to be a problem. Next year, special effort should be made to obtain a higher rate of response from the membership on nominations.

*Membership Committee and Committee on Nomination to Fellow Status.* Joe Byers, Chairman, reported that a request for nominations for Fellows in the Division was placed in the *Educational Psychologist.* To date he has received three nominations of individuals for fellowship status and 329 inquiries about membership. Nominations for membership in the Division with the status of Fellow, Member, or Associate shall be made upon the recommendation of these Committees and by not less than two-thirds vote of the Executive Committee. Candidates nominated according to these procedures will be presented at the Division Annual Business Meeting next September and declared elected to membership based upon an affirmative majority vote of those present at the meeting and the Association's approval. The APA Central Committee on Fellows must approve all Division nominees to this status. Dr. Byers will be in contact with individuals regarding their request for membership to inform them of the procedures and status of their request.

*Program Committee.* Frederick J. McDonald, Chairman of the Program Committee reported that the Committee met in the Fall to revise the Division Call for Papers. The revised call for papers appeared in the November 15th issue of the *Educational Psychologist.* This year the Division was given 38 hours for its program.

In the future, he suggested that the definition of sessions under the format of symposia be more clearly made. Papers he has received under the designation of symposium seem to fit better under the designation of related papers. He also suggested that the APA Central Office be asked to print Division regulations in their general call for papers as it is confusing for the membership to have to follow APA regulations, and separate Division regulations when they are printed in different publications at different times. The Executive Committee agreed with his suggestions and asked Dr. McDonald to explore this with Candy Won of the APA Central Office. At Dr. McDonald's suggestion, it was also decided that the program committee should have two meetings. The first meeting should be in September at the Annual Convention and should include those members of the Program Committee who were responsible for setting up the Division's Program at the on-going annual meeting and newly appointed members to the Committee. The second meeting of the Program Committee should be held to evaluate the papers which have been submitted and make final decisions regarding the Division's program.

*Committee on Thorndike Award.* Robert Glaser reported for Frederick Davis, Chairman of the Committee. The award recipient has been named and as is customary, will receive the award and make an address at the annual meeting in Hawaii.

*Ad Hoc Committees.* John Feldhusen, Chairman of the Ad Hoc Committees, reported that he will meet with Ad Hoc Committee Chairmen on Thursday, April 6.