# PLACEMENT FROM A DECISION-THEORY FRAME OF REFERENCE[1]

John R. Hills
Florida State University

You are all acquainted with the placement problem and procedures within the traditional educational framework. I have been asked to present to you some different ideas about educational placement based on decision theory, a statistical development of the last two decades. Much of what I will say is founded on the work of Lee J. Cronbach and Goldene Gleser as presented in their important book, Psychological Tests and Personnel Decisions, initially published in 1957 and revised in 1965, available from the University of Illinois Press.

Decision theory approaches the placement problem not from the position of what measures are the most accurate or the best predictors, but what procedures yield the best decisions when one allows for the expense of obtaining the data used in making the decisions. Clearly it is expensive to try to train as an airplane pilot someone who wrecks a plane while failing. Equally clearly, it would cost a lot of money to give enough tests to be 100% certain that a pupil would be successful in flight training. Decision theory evaluates decision procedures taking such costs into consideration. The best decisions are those for which the payoff, or utility, for the decision maker is the greatest after subtracting these costs.

Utility or payoff is the measure of the outcomes of various decisions. For decisions made by an institution, such as a college, one attempts to maximize the average gain or minimize the average loss over many decisions. This requires that utilities be additive--that we can express the value of various outcomes of decisions in "equal units of satisfaction" or the like which are additive.

One kind of utility scale that is handy for many purposes is that of money--dollars and cents. Units and intervals are equal on such a scale. A rational person might be happy with a set of 10 decisions that gave him five $5 bills and five times gave him nothing. He should be equally happy with a set of 10 decisions that gave him $2.50 each time. Either way he ends up with $25. This is the nature of an additive scale.

---

[1]Remarks to a session on placement at the annual meeting of the College Entrance Examination Board, Chicago, October 24, 1967.

In education, one might think of grades as being our traditional formalized utility scale. We tend to think of placing students in courses so that they won't get failing grades, and in fact, will get satisfactory grades. The grade scale is not entirely satisfactory as a utility scale for several reasons. We might not be as satisfied with 5 A's and 5 F's as with 10 C+'s. Also, ideally costs should be expressible in the same scale as utilities, and it is hard to think in terms of how many A's and how few F's would be equivalent to spending $10,000 on testing.

However, to make the rest of this presentation easier to understand, we will act as though grades were on a scale of equal intervals, and we will ignore the cost factor. (Let's act as though the students paid all the testing costs, so for the institution those costs are nil. There are some other costs, such as the waste involved in failures and drop-outs, but we will ignore them, also.) We will assume that we want a placement procedure that will maximize the average gain in grade-point-average for a group of students. Those of you who object to this utility scale can translate our discussion to whatever additive utility scale you prefer.

One of the principal characteristics of the decision-theory approach which is distinctly different from the more customary views about placement is the idea that people have different degrees of various traits, and that various traits are differently associated with success in specific learning situations. Ideally we should organize education so that there would be available a variety of ways of approaching a subject matter, and people should be assigned to the ways (treatments) according to their relevant traits so that learning would be most efficient. The fact that people with different degrees of a relevant trait learn best from different teaching procedures is called a "trait-treatment interaction."

The ways people differ might be their past experiences with the subject matter (their level of specific prerequisite training), their aptitudes, their cognitive styles, their personality traits, their attitudes, or such other characteristics as are found to be relevant to performance on the learning task. Sex might be such a variable. A golf instructor recently told me that his approach to men learners is different from the method he uses for women because the novice men tend to be rigid and inflexible with a golf club in their hands while the women tend to be loose, flexible, and wobbly. Different teaching methods are appropriate for people with different degrees of this relevant trait. This is a trait-treatment interaction.

The ways that the subject matter might be approached (treatments) might involve not only the level of prerequisite

knowledge required, but also such things as the mode of presentation, the manner and attitude of the teacher, the degree to which independent vs. group study was emphasized or provided for, the kind of teaching aids used and the extent of their use, the actual material used as subject matter, etc. When utility is expressed in the number of touchdowns made, a lightweight back might be taught an entirely different approach from a husky heavyweight who could bull through the tight places.

To get back to the ivy halls with our illustrations and to contrast the decision-theory approach with the traditional approach to placement, consider a placement procedure which might be used in English.

Suppose that the English department of a junior college complains that many of the entering freshmen do not know grammar and punctuation, have limited writing skills, and really do not know how to read fluently. Without such skills as these they simply cannot be expected to be able to write satisfactory term papers in the introductory English literature course. So a special noncredit remedial English course is proposed to develop these skills in those students who lack them, and a placement test is provided to determine which students should be assigned to the remedial course.

Now, according to the traditional model, the students who are assigned to what will probably soon be dubbed "bonehead English" and who pass that course, should then be expected to be able to compete in the standard freshman introductory English literature course on a par with the other incoming freshmen. Their "holes" have been patched, so to speak. Statistically, if the remedial program is working soundly we should not be able to distinguish between the grade distributions of introductory English literature students who have taken and passed remedial English and those who were not assigned to remedial English but went directly into the regular freshman course.

In terms of the decision-theory model, we have two treatments in connection with freshman English. One is two courses long, consisting of remedial English followed by regular freshman English. The other is one course long, that single course being regular freshman English. We also have a placement test to determine which students will be assigned to each of the treatments.

Now, suppose that a student who has mastered the skills in grammar, punctuation, writing, and reading will not be benefited by taking the two-course treatment. Further, suppose that he may be so bored by that treatment that he will do less well in the second part of the sequence, the regular freshman English literature part which is our

basic interest, than he would have done if he entered that course directly. Suppose, on the other hand, that the student who has not mastered the basic English skills will be very handicapped by the one-course treatment, but be given a great advantage in regular freshman English by the two-course treatment. This is a trait-treatment interaction.

The situation may be illustrated by Figure 1. The lines in the figure are the regression lines, or prediction lines, of grades in regular freshman English literature on placement test scores. It can be seen that if a student's score on the placement test is to the right of the crossing point, he can be expected to obtain higher grades in the regular literature course if he is placed in the one-course sequence. It might be that he merely avoids boredom. However, if his placement test score is below the crossing point, he will obtain higher grades in the regular English part of the instruction if he takes the two-course treatment. In this case, the student's difficulties may have been corrected before facing the regular course.

Notice the striking thing that this decision-theory model says about our placement test. For placement to be worthwhile, the placement test must have different regression slopes for the various treatments. The traditional placement testing model does not recognize this necessity, and, indeed, traditional practice in colleges fails to recognize it also. I cannot recall ever seeing a claim that a commercial or locally-developed placement test was efficacious because it had different regression coefficients for different available treatments. A test which measures ability in general, or which correlates highly with a general ability test, will be of little value in placement under this model because very similar regression slopes will result for all treatments. The interaction between trait and treatment is the key to the model. In educational terms, such an interaction is no more than recognizing and capitalizing on the fact that different pupils can best be taught by different methods.

This approach suggests that a college should seek, or attempt to create, trait-treatment interactions involving the subject matters wherein placement is a concern. For instance, in a subject-matter such as mathematics, English, or a foreign language, one might attempt to invent several different approaches in accordance with his hypotheses as to the bases of the difficulties common to different types of students. He then might attempt to develop a test which would have different regression slopes for the different treatments. Students would be assigned to the treatments according to the locations of their scores on the test in relationship to the crossing point of the regression lines. It might or might not be the case that one of the effective

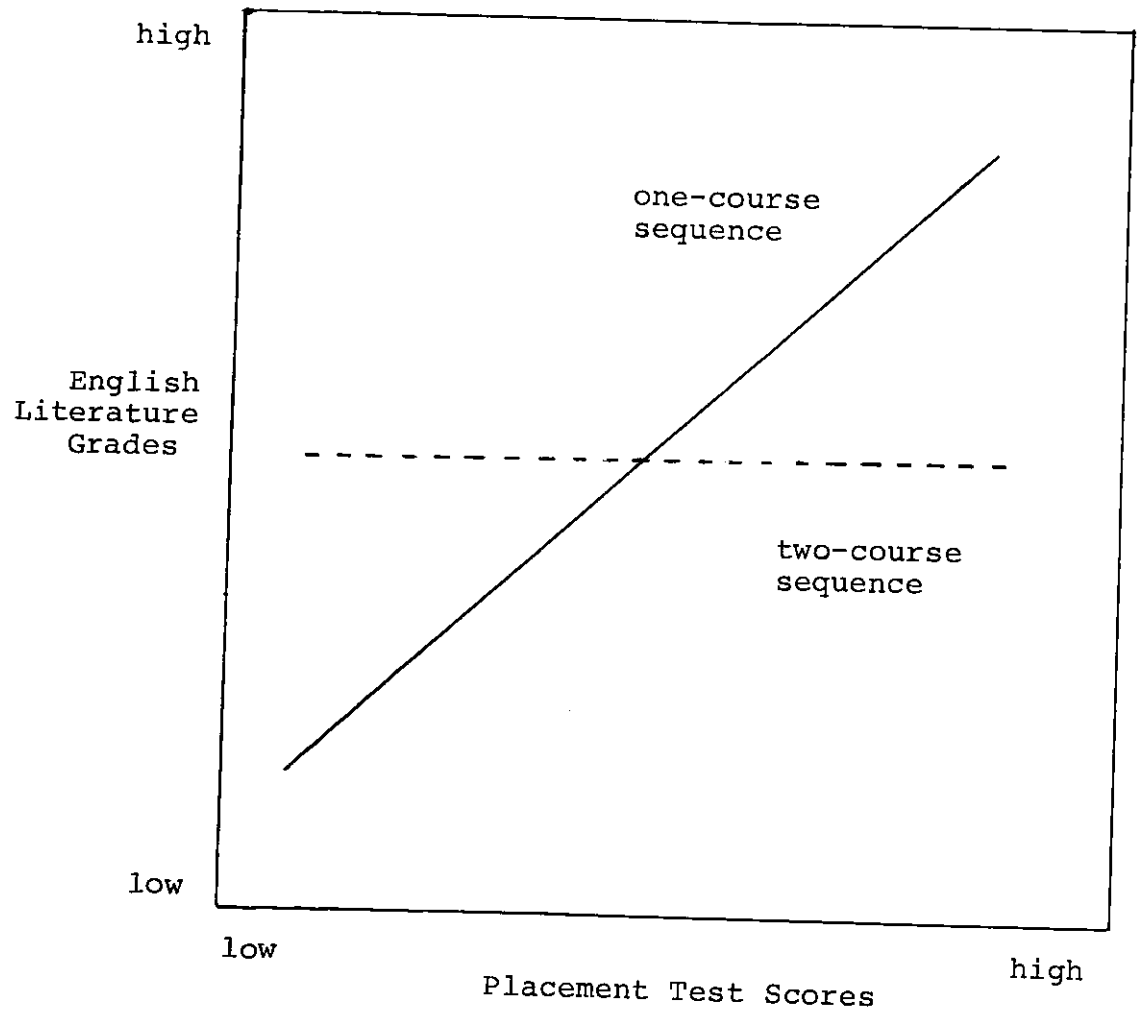


Figure 1. Regressions of English Literature Grades on Placement Test Scores Where Placement is Effective

treatments would be the same as what is now done under the label of remediation or advanced placement in that particular subject matter.

Some of the problems in using the decision-theory approach are already discernible. As the college about which we have just been talking attempts to find or create tests with treatment interactions, it will probably discover that the "placement tests" that are now available commercially will not be relevant for placement based on interactions. Those tests were not designed with different treatments in mind, and they were validated, if at all, against course grades or test scores without regard to the kind of instruction that produced those grades or scores.

It may be that we can never expect commercial publishers to produce tests suitable for placement based on trait-treatment interactions. The number of kinds of treatments that are discovered and used may be so great that the trait measures will also be very heterogeneous and thus generate little sales volume. However, some kinds of trait-treatment combinations may be found to be generally useful. Their widespread adoption would create the demand for a commercial test product that could be profitable. In fact, if such combinations are found, the commercially satisfactory development may be an entire trait-treatment package, including the measure or measures of relevant traits, preselected cutoff scores, the method of teaching (perhaps in a computer program or other programmed materials), and the evaluative criterion measure.

If our college gives up on finding appropriate tests, it may be reduced to developing its own. This will require measurement sophistication of a high order. The unsophisticated test constructor will probably tend to produce a general-ability measure which will fail to have the necessary interactions with different treatments. Beyond that, there are severe reliability requirements on placement measures. If one trait is considered in relation to several treatments, the standard error of measurement of that trait should be as small as possible so that as few students as possible obtain scores near the cutting point. Peaked tests might be needed after cutting points have been determined. Further, internal-consistency reliability, while necessary, will not be sufficient. The traits and their measures must also be stable over substantial periods of time--at least as long as is required to process the students through the related learning sequence.

Probably our college will look to its faculty and staff to develop its new placement tests. It remains to be seen to what extent teachers will be willing and able to work with researchers to take full advantage of trait-treatment interactions. Some faculty members will probably

recognize the significance of this approach and be excited about doing something constructive with it. It seems likely that many individual faculty members will resist the idea because of its novelty and because it will tax their knowledge and understanding of their own subject matter, of teaching procedures, and of the nature of their fellow men. The faculty may have difficulty even suggesting variables which can be expected to interact with various instructional procedures, and they may not be able to describe their own teaching procedures, i.e., treatments, well enough for anyone else to generate useful hypotheses about interacting variables.

Other faculty problems can also be foreseen. It may be that the interacting variables are such that the usual modifications of classroom procedures for placement are not relevant, i.e., watering down or beefing up the content, shortening or lengthening the treatment through added or deleted explanations, drill, examples, etc. Unorthodox modifications may be necessary, expecially if the interacting variables are in the personality domain. It may be the teacher's attitude which must be modified, or the classroom structure, or entirely new approaches may be required for some students. It may prove necessary to replace some faculty with others who hold different attitudes instead of asking the old guard to abandon their entrenched ways. Even if all that is done, work by some of my FSU colleagues suggests that it will probably be difficult to invent long sequences of instruction based on specific instructional procedures. Probably the practice in the past has been for teachers to vary their treatments, intuitively recognizing interactions between traits and treatments and taking them into account by providing a variety of treatments so that all traits would be accommodated. If students were grouped according to known trait-treatment interactions, it might be that a lot of time would be saved by not using all the varied teaching techniques on everyone.

This leads to another set of problems that can be foreseen. If use of trait-treatment interactions does speed learning and make it more efficient, scheduling problems will develop for the administration. In the past the administrative problem was as simple as deciding how many faculty could be assigned to remedial or advanced work and assigning classroom space, letting the chips fall where they would. Now suppose that it is discovered that one of the principal problems with certain students in mathematics is a fear of numbers. Further, suppose that a measure that would detect and evaluate this antipathy is developed, and that a treatment is developed which would reduce or eliminate it in a matter of six weeks of concentrated effort. The administration would have to figure out what

to do with such students at the end of six weeks when they were all ready to go into college mathematics, but there would not be a mathematics class ready for them until the next term started.

This is one kind of flexibility that will have to be created. Suppose that instead of only one kind of remediation in mathematics it turns out that there are really three or four different kinds of maladies in mathematics, each requiring a different kind of treatment. Provision for each treatment would probably have to be made available several times a day in order for each student to have an opportunity to schedule his proper treatment--or treatments. Suppose further that several special treatments are to be provided in each of a variety of subjects. Scheduling soon becomes terribly complex--but there are already data available that suggest that if this problem is ignored or if proper assignment is only roughly approximated, placement on the basis of trait-treatment interactions yields inconsequential results.

Another administrative problem will be the matter of quotas. For effective placement, quotas may have to be adaptable, set according to the students who are available to be placed. For the administration to have to become adaptive to providing a set of programs and letting the students get the best education they can out of it, will require extensive adjustment in many cases. To some extent the adaptiveness that would be optimal may not be possible, and it may be that the compromises necessary for administrative feasibility will destroy the benefit to be expected from placement. This may, indeed, be partly what has resulted in placement as it is now practiced having so little to show for itself.