DOCUMENT RESUME

ED 394 458 HE 029 154

AUTHOR	Sawyer, Richard
TITLE	Assessing the Effectiveness of Course Placement Systems in College.
PUB DATE	10 Apr 96
NOTE	65p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (New York, NY, April 10, 1996).
PUB TYPE	Speeches/Conference Papers (150) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Admission (School); Educational Diagnosis; *Educational Testing; Higher Education; *Placement; *Remedial Mathematics; Remedial Programs; Screening Tests; Student Evaluation; *Testing
IDENTIFIERS	*Decision Theory; *Entry Level Mathematics Examination

ABSTRACT

This paper illustrates how statistical decision theory can be used to model aspects of the effectiveness of a course placement system. The illustration is based on data from students who enrolled in a first-year mathematics course at a midwestern university. To model the placement of students, the author first elicited students' and instructors' preferences for the different outcomes of the course. Completed responses were received from 19 instructors and 40 percent of the 270 students questioned. About 3/4 of the student respondents and 4/5 of the instructor respondents were able to supply coherent grade value functions; the median value functions for the instructors were modestly but systematically higher than those for the students. Comparisons were then made of the actual outcomes of students who received remedial instruction before enrolling in the course with their predicted outcomes had they not done so; an analogous comparison was made between expected value functions. Although the number of students with previous remedial instruction was too small to permit drawing firm conclusions, results illustrate how, given sufficient sample sizes, one could use decision theory to develop indicators of the effectiveness of remedial instruction. Appendixes include the questionnaires and discussion of constructing summary value functions for course placement outcomes. (Contains 36 references.) (Author/NAV)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 394 458

Assessing the Effectiveness of Course Placement Systems in College

Richard Sawyer

Research Division,
American College Testing
P.O. Box 168
Iowa City, Iowa 52243

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it
Minor changes have been made to improve reproduction quality
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Richard Sawyer
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Paper presented at the annual meeting of the
National Council on Measurement in Education, New York

April 10, 1996

TM025154

BEST COPY AVAILABLE

Abstract

Course placement systems in college include both an assessment component (to estimate students' probability of success in standard first-year courses), and an instructional component (in which underprepared students are taught the skills and knowledge they need to succeed in the standard courses). The effectiveness of a placement system depends on students' eventual success in the standard courses. Success is usually defined in terms of completing the standard courses with satisfactory grades.

This paper illustrates how statistical decision theory can be used to model aspects of the effectiveness of a course placement system. The illustration is based on data from students who enrolled in a first-year mathematics course at a midwestern university. To model the placement of students, I first elicited students' and instructors' preferences for the different outcomes of the course. I then compared the actual outcomes of students who received prior remedial instruction before enrolling in the course with their predicted outcomes, had they not done so. I made an analogous comparison between expected 'value functions'. In this particular example, the number of students with previous remedial instruction was too small to permit drawing firm conclusions. Nevertheless, the results illustrate how, given sufficient sample sizes, one could use decision theory to develop indicators of the effectiveness of remedial instruction.

Assessing the Effectiveness of Course Placement Systems in College[1]

Richard Sawyer

A typical and important use of college entrance tests is course placement, i.e., matching students with instruction appropriate to their academic preparation. For example, students whose academic skills are insufficient for them to be successful in a standard first-year English course might, on the basis of their test scores and other characteristics, be advised or required to enroll in a remedial English course. On the other hand, students with an unusually high level of academic preparation might be encouraged to enroll in an accelerated course or in a higher-level course.

Most colleges and universities enroll students who are not academically prepared to do work at a level traditionally expected of their first-year students. The percentage of postsecondary institutions with some form of placement and remedial instruction has steadily increased in the past decade, and is now about 90% (Woods, 1985; Wright and Cahalan, 1985; McNabb, 1990; "Colleges and Universities Offering Remedial Instruction," 1994). A recent survey by the American Council on Education (1996) suggests that about 17 percent of students in community colleges and about 11 percent of students in public four-year institutions take remedial courses. One suggested explanation for the growth of remedial instruction is that American high schools have become less effective in preparing students for college (The National Commission on Excellence in Education, 1983; The Carnegie Foundation for the Advancement of Teaching, 1988; Singal, 1991). Another explanation is that more students from disadvantaged backgrounds are attending college (Munday, 1976; College Entrance Examination Board, 1977; Carriuolo, 1994).

During the past three decades, several authors have proposed using statistical decision theory to validate educational selection systems. Cronbach and Gleser (1965) adapted linear regression methodology to estimate the expected costs and benefits of using a test score or other predictor variable for classifying or selecting personnel. Their technique continues to be widely applied in employment selection. Petersen and Novick (1976) developed a model based on Bayesian decision theory. Ben-Shakhar, Kiderman, and Beller (1994) compared these two approaches, and illustrated them using data from an admission selection problem.

[1] I am grateful to Dan Anderson and to Jerry Dallam for their help in collecting data for this study.

Sawyer (in press) proposed a statistical decision theory model for validating course placement variables such as tests. The model can be used to compare the effectiveness of alternative placement variables in identifying underprepared students, and to determine appropriate cutoff scores on these placement variables. Sawyer (1994, 1995) proposed another, more comprehensive decision theory model for measuring the effectiveness of remedial instruction, and he studied alternatives for eliciting preferences for the possible outcomes. The present paper applies the more comprehensive model to a first-year mathematics course at a midwestern university. Elicited preferences for outcomes are combined with actual course outcome data to make inferences about the benefit of prior remedial instruction for subsequent academic achievement.

Background

Remedial Instruction

At many postsecondary institutions, there are two levels of first-year courses: a 'standard' course in which most students enroll, and a 'remedial' course for students who are not academically prepared for the standard course. At some institutions, the lower-level course may be given other names, such as 'college-preparatory', 'compensatory', 'developmental', or 'review'. Carriuolo (1994) articulated differences in the meanings of 'remedial' and 'developmental'. Often, remedial courses do not carry credit toward satisfying degree requirements. At some institutions, there may be courses that require more knowledge and skills than the lowest-level remedial course, but less than the standard course. In the model considered here, only a single lower-level course is considered, and it is designated 'remedial', to be consistent with Willingham's (1974) nomenclature.

Though essential to placement, testing is but one component of an overall system. To be educationally effective, a placement system must satisfy all of the following requirements:

1. Students who have small chance of succeeding in the standard course (underprepared students) are accurately identified.
2. Appropriate remedial instruction is provided to these underprepared students.

3. Both the students who originally enrolled in the standard course, and the students who were provided remedial instruction, eventually do satisfactory work in the standard course.

Note that accurately identifying underprepared students (Requirement 1) is necessary, but not sufficient, for a placement system as a whole to be effective. Accurate identification is not an end, but only a mechanism for effectively allocating remedial instruction (Requirement 2). On the other hand, providing remedial instruction is itself only a means to achieve the larger goal that students succeed in college: Even if underprepared students are accurately identified and are provided remedial instruction, if they eventually drop out or fail in the standard course, then little will have been accomplished by the placement system. On the contrary, both the institution's and the students' resources will have been wasted. Van der Linden (1991) noted that a defining characteristic of course placement systems is that students take different treatments (courses), and the success of each treatment is measured by the *same* criterion variable.

One might argue that failure in the standard course can lead to positive results for students, such as their selecting and succeeding in another educational program better matched with their talents and interests. While this statement is undoubtedly true for some students, they would have done better to select their preferred educational programs in the first place, through appropriate counseling. This scenario illustrates that effective counseling is important for effective placement. This paper does not, however, attempt to model the effect of counseling on the outcomes of placement.

The need for an institution to serve students who by traditional standards are academically unprepared for college imposes a fourth requirement on placement systems. Even if a large proportion of the underprepared students are accurately identified, are provided remedial instruction, and ultimately succeed in the standard course, the overall result still might not be satisfactory. This would occur if an institution diverted resources to instruction in the remedial course to such an extent that the performance of students in the standard course was adversely affected. In other words, institutions should consider the tradeoffs they must make in allocating their finite resources when they provide remedial placement

systems; such considerations may relate to institutional mission and policy, as much as to costs and to grades. A related concern is that publicly-supported educational institutions must compete for money with agencies that serve other pressing social needs. There is controversy about the proper role of remedial placement in the missions of postsecondary institutions, particularly in publicly-supported institutions. Mac Donald (1994) argued that by overexpanding its remedial programs, the CUNY system seriously degraded the quality of its standard-level undergraduate programs. Lively (1993) reported on efforts in different states to eliminate remedial instruction from four-year public institutions by designating that role to two-year colleges. The American Council on Education (1996) reported on similar efforts in Congress.

Non-cognitive variables in course placement systems

Academic achievement, as measured by persistence and grades in standard courses, is certainly an important outcome of course placement, but it is not the only important outcome. For example, most colleges want their students to achieve other, non-cognitive goals (e.g., working well in groups, understanding and respecting different points of view, developing career and intellectual goals). Furthermore, certain non-cognitive factors (e.g., prior work experience; hours worked on a job while in school; ethnic background) are well-known to be important factors in academic success. For both these reasons, effective course placement requires effective counseling (ideally, one-on-one), and *vice versa*.

Non-cognitive variables are also important in evaluating the effectiveness of a course placement system. For example, do students believe the advice they have been given was appropriate? Do students think that they have been treated well by 'the system'? Do the faculty and staff who run the system believe that their skills are effectively used, and that *their* needs are considered? While issues such as these are beyond the scope of this paper, they need to be addressed in evaluating the overall effectiveness of course placement systems (Frisbie, 1982).

A Decision Theory Model for Course Placement

The decision problem can be formally defined as follows: One must select a particular decision d from a set D of possible decisions. A particular outcome θ occurs, from among a set of possible outcomes Θ. A utility function $u(d,\theta)$ assigns a numerical value to the desirability of decision d when the outcome is θ. The exact outcome θ that occurs is unknown to the decision maker, but there is some probabilistic information available about the likely values of θ. In a Bayesian decision model, this information is described by a subjective probability distribution on Θ; the subjective probability distribution quantifies the decision maker's personal beliefs about the likely values of θ, given both prior beliefs and any relevant data collected. The Bayesian optimal strategy is to choose the decision d that maximizes the expected value of $u(d,\theta)$ with respect to the subjective probability distribution on Θ (Lindley, 1972). (See p. 8 for a discussion of expected utility functions.)

To apply this tool to course placement, suppose there is a cutoff score K on a placement test, and that:

* test scores are obtained for all first-year students at an institution;
* students whose test scores are less than K are provided remedial instruction before they enroll in the standard course, and students whose test scores are greater than or equal to K enroll directly in the standard course; and
* the actual final performance in the standard course is known for all students (i.e., for students who are provided remedial instruction, as well as for those who are not).

The final performance in the standard course of students who first enroll in the remedial course will, of course, become known later than the performance of students who enroll directly in the standard course. For each student, four possible events could occur, as shown in Table 1 on the following page.

Table 1
Events Associated with Identifying and Providing Remedial Instruction to Underprepared Students

Event	Test score	Course into which student is placed	Eventual performance in standard course
(1)	≥ K	Standard	Successful
(2)	≥ K	Standard	Unsuccessful
(3)	< K	Remedial	Unsuccessful
(4)	< K	Remedial	Successful

Each student is classified either as being adequately prepared for the standard course (if her or his test score equals or exceeds the cutoff score K), or as needing remedial instruction (if the score is less than K). Because the classification for any student depends on K, the set of decisions (*D*) in this case is the set of possible values of K. The goal is to find the 'best' value of K, and to quantify the effectiveness of the associated instruction.

At an institution without a placement system, the events in Table 1 could be observed as follows:

* Randomly assign students, regardless of their test scores, either to enroll directly in the standard course or to enroll first in the remedial course.
* Observe the students' eventual performance in the standard course, and note which of them succeed and which do not succeed.

For each value of K, there would be a set of proportions associated with the events (1) - (4). Let us suppose, temporarily, that data are collected this way; the modifications required when there is prior selection resulting from an existing placement system are described on p. 10.

Let $p_1(K)$, $p_2(K)$, etc., denote the observed proportions of students corresponding to events (1), (2), etc., in the entire group of students when the cutoff score is K. Then, for example, $p_1(K) + p_4(K)$ is the proportion of students who are ultimately successful, and $p_2(K) + p_3(K)$ is the proportion of students who

are ultimately unsuccessful. The overall usefulness of the predictions can then be evaluated in terms of the costs and benefits associated with each event (1) - (4). A function that assigns a value to outcomes such as these is called a utility function. One class of utility functions would assign different values to each event, and weight their sum:

$$u(K;\theta) = w_1p_1(K) + w_2p_2(K) + w_3p_3(K) + w_4p_4(K) \quad (1)$$

where $0 \le w_1,\ldots,w_4 \le 1$. Such a function would quantify the different costs and benefits of each outcome. Consider, for example, the trade-offs a student must make in his or her utility. Although students pay tuition to take remedial courses (just as they do to take other courses), remedial courses often do not carry college credit. From a student's perspective, the weights $w_1, \ldots, w_4$ must balance the likely improvement in her or his eventual performance in the standard course against the extra time and money spent on taking the remedial course.

In principle, utility functions are person-specific, and hence need to be elicited separately for each student, counselor, teacher, or administrator. In practice, this is not feasible, and we must look for utility functions that reasonably approximate the preferences of different groups of people.

Other Models

In the model described in Table 1, there are only two results in the standard course: 'Successful' and 'Not successful'. In practice, 'Successful' usually means completing the standard course with a particular grade (e.g., C) or higher. The adequacy of the model in Table 1 therefore assumes that the decision maker's preferences for particular grades have a step-function relationship. Petersen and Novick (1976) called such a function a 'threshold utility'. Results obtained by Sawyer (1995, 1996) suggest that in course placement, most students' and instructors' preferences are not well described by the threshold utility.

A more complex decision theory model, defined directly in terms of the grade received, would describe people's preferences more accurately. For example, instead of designating each student as 'Successful' or 'Unsuccessful' in the standard course, one could specify the student's completion of the

course and final grade (e.g., A-F). In this case, there would be 10 outcomes (rather than 4) in the model; such a model is described on p. 20. If we also considered withdrawal before completing the standard course as a separate kind of unsuccessful outcome, then there would be 12 outcomes in the model.

Expected Utility Functions

In practice, a utility function cannot be directly computed for the group of students for whom placement decisions are to be made, because the actual outcomes (students' test scores and eventual performance in the standard course) are not yet known. In (1), for example, the actual proportions $p_1(K)$, $p_2(K)$ etc., are not known for a particular group of students before they are tested and complete the standard course. These proportions must instead be estimated in some way from data on past students, under the assumption that future students will be similar to past students.

The 'expected utility function' is a formal mechanism for dealing with the uncertainty of outcomes in a decision theory model. It is from the expected utility function that decisions on the effectiveness of a placement system can be made. In Bayesian models, an 'expected utility' for a decision *d* is the average (expected) value of a utility function $u(d,\theta)$ with respect to the decision maker's subjective probability distribution for the outcomes Θ. In the example previously given,

$$u'(K) = E_\theta[\ u(K,\theta)\ ] = w_1\tilde{p}_1(K) + w_2\tilde{p}_2(K) + w_3\tilde{p}_3(K) + w_4\tilde{p}_4(K) \qquad (2)$$

where $\tilde{p}_1(K) = E_\theta[\ p_1(K)\ ]$, $\tilde{p}_2(K) = E_\theta[\ p_2(K)\ ]$, etc., are estimated from a past group of students. In the Bayesian model, the estimates $\tilde{p}_1(K)$, $\tilde{p}_2(K)$, etc., are the expected values of the corresponding observed proportions with respect to the decision maker's subjective probability distribution for students' test scores and course grades. In the terminology of Bayesian statistical inference, the subjective probability distribution for test scores and course grades is specified by a 'predictive density' for their joint distribution. The predictive density is based on prior beliefs about the joint distribution and on data obtained from a particular group of past students. Although simple in concept, Bayesian statistical methods can be mathematically formidable in real applications. When prior beliefs are vague or as sample sizes become large, however, Bayesian estimates are, for practical purposes, similar to much simpler

ERIC

estimates based on classical sampling theory (i.e., estimates based only on an assumed model and on data; DeGroot, 1970).

Sawyer (in press) described a simple procedure, based on sampling theory, for estimating the cell probabilities $p_1(K)$, $p_2(K)$, etc. The first step is to estimate the relationship between success in the standard course and a placement test score using a logistic regression function:

$$P[Y=1 \mid X=x] = (1 + e^{-\alpha-\beta x})^{-1} \tag{3}$$

where Y = 1, if a student is successful,

= 0, if a student is unsuccessful;

and X is the student's score on a placement test or other placement variable. The numbers α and β in Equation (3) are unknown parameters that are estimated from data on the test scores and on the success/failure variable Y for a group of enrolled students. The regression function $P_S(x)$ of students who enroll directly in the standard course and the regression function $P_R(x)$ of students who enroll first in the remedial course are estimated separately.

Once estimates *a* and *b* have been obtained for the unknown parameters α and β, the conditional probability of success can be estimated by substituting *a* and *b* in Equation (3). From the estimated conditional probabilities $\hat{P}_S(x)$ and $\hat{P}_R(x)$, expected utilities can be calculated. In the threshold model described by Table 1, for example, the proportion of students associated with Event (1) in Table 1 can be estimated by:

$$\hat{p}_1(K) = \sum_{x \geq K} \hat{P}_S(x) * n(x) / N \tag{4}$$

where $\hat{P}_S(x)$ = estimated P [Y = 1 | X = x] for students who enrolled directly in the standard course,

K = the minimum score required for enrollment in the standard course (cutoff score),

n(x) = the number of students in the placement group whose test score is equal to x, and

N = $\sum_{\text{all } x} n(x)$, the total number of students in the placement group.

The proportions for Events (2), (3), and (4) can be estimated similarly, using the appropriate conditional probability function $\hat{P}_S(x)$ or $\hat{P}_R(x)$, and the appropriate region $[x \geq K]$ or $[x < K]$.

The procedure can readily be extended to more complex models. In the 10-outcome model (A-F/remedial instruction or not) described on p.7, for example, one would first estimate the conditional probabilities of the 10 outcomes y, given the test score X. Then, one would sum, over the outcomes y and over the corresponding appropriate values of x, the products $u(y) * \hat{P}(y|x) * n(x) / N$. An example illustrating such computations is given on pp. 19-20.

Effects of Prior Selection

Note that the summations in Equation (4) are based on the x-values (e.g., test scores) of all the students in the placement group (the set of students for whom placement decisions are made), not just the students who complete the course. In practical terms, the placement group usually consists of all first-time entering students with test scores, regardless of which course they actually enroll in. Of course, one could also define a placement group for students in a particular program of study (e.g., business) or with particular background characteristics (e.g., minority students).

At an institution with an operational placement system with cutoff score K_0, we can estimate $P_S(x)$ only from data with $x \geq K_0$, and we can estimate $P_R(x)$ only from data with $x < K_0$. The reason is that students whose test scores are below the cutoff score K_0 do not enroll directly in the standard course, and therefore do not have performance data unaffected by remedial instruction. Sawyer (1993) noted, however, that the logistic regression model (3) can be conveniently extrapolated to test scores below the current cutoff score K_0. Schiel & Noble (1993) compared logistic regression functions estimated from truncated subsets of a data set that was not subject to prior selection. They found that when the truncation involved less than 15% of the population, the resulting errors were small, but that large amounts of truncation (e.g., 50%) resulted in large errors. Houston (1993) did computer simulations to examine the effects of truncation on the accuracy of estimated conditional probabilities of success. He found increases in standard error of 6%, 30%, and 43% when the placement group was truncated at the

25th, 50th, and 75th percentiles, respectively, as compared to the standard error associated with no truncation.

Optimal Cutoff Scores

If the expected utility $u'(K) = E_\theta[u(K;\theta)]$ attains a maximum value at some cutoff score K_0, then using K_0 as a cutoff score will result in a greater expected utility for the group than using any other cutoff score. Furthermore, if K_0 is between the minimum and maximum possible scores on the test or other placement variable, then the effectiveness of the placement system as a whole is supported. On the other hand, if u' is an increasing function, then the effectiveness of the placement variable is called into question --- the placement variable is not able to discriminate between students who should enroll directly in the standard course and those who should first take the remedial course. Finally, if u' is a decreasing function, then the effectiveness of both the placement variable and the remedial course is called into question. Of course, all of these inferences depend on the validity of the success criterion variable.

Eliciting Utility Functions

If the decision model is to be useful in real applications, its utility function must accurately describe the preferences of the decision makers. In the model described by Table 1, for example, we need some way to quantify students' and instructors' preferences for success in the standard course, as balanced against the extra time and cost associated with taking the remedial course.

There is a vast literature on eliciting (i.e., assessing) utility functions. One important characteristic distinguishing various utility theories is whether they are deterministic or stochastic:

* A *value function* measures the satisfaction of any sort of 'want' without regard to uncertainty. For example, some economists model the satisfaction that an individual receives from consuming commodities. The key characteristic of a value function is that it assigns numerical values to the subjective worth of outcomes without regard to uncertainty (Yates, 1990).

  A simple example of eliciting a value function would be to ask an individual to rank each possible outcome on the following Likert scale:

1='dislike very much', 2='dislike', 3='dislike a little', ..., 7='like very much')

Note that in this example, the assignment of values to outcomes is done outside any context of uncertainty or risk.

* A *von Neumann-Morgenstern* utility, in contrast, is explicitly defined in terms of uncertainty. The standard assumption in von Neumann-Morgenstern (abbreviated hereafter as vN-M) theories is that the decision maker has a preference relation $\prec$ over the set Π of *probability distributions* on the outcome space Θ (rather than on Θ itself), and that $\prec$ satisfies an appropriate set of axioms (e.g., transitivity). Then it can be shown that there exists a real function u on Θ, such that for distributions $p, q \in \Pi$, $p \prec q$ if, and only if, $E_p[u] < E_q[u]$. The function u is unique up to positive, linear transformations; therefore, one can without loss of generality assign the value 0 to the least favorable outcome and the value 1 to the most favorable outcome. Note that vN-M utility functions are defined in terms of probability; therefore, their elicitation is naturally done in reference to hypothesized probability distributions (typically, hypothetical lotteries). See Farquhar (1984) for a comprehensive review of different strategies for eliciting vN-M utility functions.

The principal advantage of value functions is that they are easy to elicit, because they do not require any reference to uncertainty or risk. The principal advantage claimed for vN-M utility functions is that they are more realistic, because they reflect the decision maker's feelings about both the inherent worth of the outcomes, and the risk involved in making choices. On the other hand, this realism is elicited in the context of hypothetical situations. Although both value functions and vN-M utility functions can formally be used in expected utility models (Yates, 1990), they are not the same, and can lead to different decisions. I shall follow Yates' convention in reserving the term 'utility' to refer specifically to a vN-M utility function, and the term 'value function' for a function that does not consider risk.

Sawyer (1996) elicited vN-M utility functions and value functions from students enrolled in either of two particular postsecondary institutions in the midwest, using a paper-and-pencil questionnaire. He

found that about half of the university students and about two-thirds of the community college students surveyed completed the questionnaire. Of students and instructors who completed the questionnaire, about 2/3 provided enough information to develop an internally consistent grade value function. For the typical respondent, only 35% of the possible vN-M utility functions could be calculated, however. These results suggest that institutions can not realistically expect to elicit utilities by administering paper-and-pencil questionnaires to their instructors and students (see Question 1, p. 16). Institutions would instead need to provide special instruction and motivation to elicit valid utilities for most students and instructors. An interactive computer elicitation program might increase validity by giving respondents an opportunity to correct inconsistencies. The difficulty and expense of implementing such a computer program, however, would seem to make it unattractive to institutions, even today.

Sawyer (1995) also found that there was a small, but consistent difference between instructors and students in the values of their elicited value functions: Instructors tended to place somewhat greater value on the grades B, C, and D (by about .05 - .10) than did the students. This result is consistent with Whitney's (1989) observation that students and teachers have different interests in course placement.

Data

Eliciting Value Functions

The value function data for this study came from students enrolled in five elementary mathematics courses at a public postsecondary institution ('Midwestern University') in fall 1995. An individual professor supervised 18 graduate student instructors of the courses. The professor and the instructors agreed to distribute and to collect questionnaires from their students. Part 1 of the questionnaire asked for background information about students' previous course work, their majors, and their goals. Part 2 of the questionnaire elicited a value function for grades in a hypothetical standard course. Part 3 of the questionnaire elicited students' preferences for taking a remedial course before taking the standard course. The professor and the instructors also completed an instructor's version of the questionnaire that collected similar information.

Respondents indicated their relative preferences for the grades B, C, and D in the standard course by marking these letters on a number scale. The number scale ranged from 0 to 100, in units of 10. The grades A and F were anchored at 0 and 100, respectively.

Respondents expressed their preferences for taking remedial courses by responding to choice tasks (e.g., "Would you prefer to earn an A in the standard course after taking the remedial course, or would you prefer to earn a B in the remedial course if you did not have to take the remedial course?"). I asked the respondents to make 10 separate choices, corresponding to the 10 possible pairings of a given grade with the grades lower than it.

Appendix A contains the two survey instruments. The documents in Appendix A have been modified to conceal the identity of the university and the catalog numbers of the courses.

Although 270 students completed the questionnaire, the overall response rate was disappointingly low, about 40%. The likely causes for the low response rate were that it was voluntary, that no remuneration was given to the students, and that it was completed at home rather than in class. The response rate did vary significantly among courses (see Table 2 on the following page). According to the professor in charge, the variation in response rates was likely caused by the extent to which instructors emphasized to students the importance of completing the questionnaire. In the future, I will reduce the length of the questionnaire so that it will be feasible to administer it in class.

The response rate for the 19 instructors was 100%.

Table 2.
Response Rates to Value Function Elicitation Survey, by Group

Group	Number of completed questionnaires	Response rate
Students		
Course 1: Basic Algebra I	19	.68
Course 2: Basic Algebra II	81	.31
Course 3: Trigonometry	52	.32
Course 4: Elem. functions	109	.43
Unknown course	9	...
Total, all students	270	.40
Instructors		
Total, all instructors	19	1.00

Summary Value Function

For each respondent, I computed a summary value function for the 10 outcomes described on p. 8. The summary value function measures a respondent's preferences for taking the remedial course or not, relative to different grades in the standard course. I computed the summary value function by combining the information elicited from Part 3 of the questionnaire with the grade value function information elicited from Part 2. The following example illustrates how the two sources of information were combined. Let 'x ≻ y' denote that "Outcome x is preferred to outcome y"; let (R,A) denote the outcome ["Student takes the remedial course before taking the standard course and earns an A."]; and let (S,B) denote the outcome ["Student takes the standard course directly and earns a B."]. Suppose that a respondent's choices in Part 3 of the questionnaire indicate that (R,A) ≻ (S,B). Then, a summary value function *svf* can be imputed by interpolating between 1=*gvf*(A) and *gvf*(B): *svf*(R,A)=.50 + .50**gvf*(B).

It is possible to do such imputation consistently provided the respondent's choices in Part 4 are coherent (consistent). For example, the following two choices are incoherent:

a. (R,A) $\succ$ (S,B)
b. (S,C) $\succ$ (R,A).

Of the 2^{10} = 1024 possible sequences of choices, only 14 are coherent. See Appendix B for details on the coherent choice sequences.

Course outcomes

Figure 1 on the following page shows the structure of placement and pre-requisites for first-year mathematics courses at Midwestern University. As is readily apparent, the structure is complex. Students can enroll in courses on the basis of their high school course work and their scores on a locally-developed and administered placement test. Students who enroll in particular lower-level courses can subsequently enroll in more advanced courses, as shown by the dotted flow lines in the diagram. For example, students who enroll in Course 4 (Elementary Functions) may do so in four different ways:

a. Directly, on the basis of their high school course work and placement test scores.

b. Initially in Course 3, then in Course 4.

c. Initially in Course 2, then in Course 3, then in Course 4.

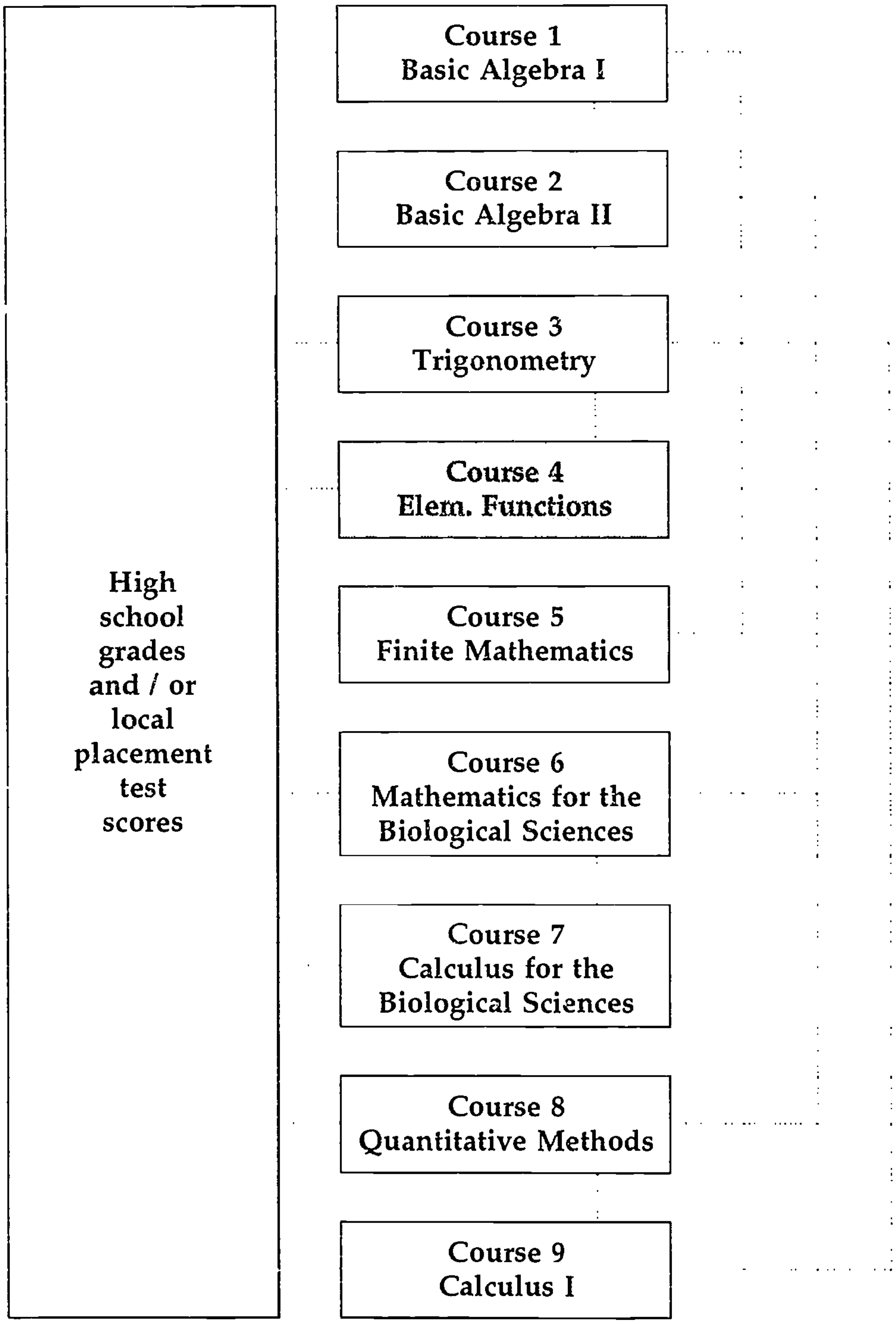
d. Initially in Course 1, then in Course 2, then in Course 3, then in Course 4.

For other courses, such as Course 9, there are many more entry paths.

The enrollments in Course 1 was small. Courses 2 , 3, and 4 had substantially more students (e.g., more than 100). Course 4 (Elementary Functions) was the lowest-level course with a substantial enrollment, and where substantial numbers of students enrolled in lower-level courses. Course 4 also has fewer ways in which students could have been previously instructed than more advanced courses with large enrollments. For these reasons, I chose Course 4 as the 'standard course' for analysis.

ERIC

Figure 1.
Structure of Course Placement and Prerequisites
for First-Year Mathematics Courses at Midwestern University



Note: Course 4 (shaded box) was taken to be the 'standard course' in the decision model.

I obtained from administrators at Midwestern University the ACT Mathematics scores (ACT, in press), course enrollment histories, and course grade data for all students who enrolled in Courses 1, 2, 3, or 4 during the years 1993, 1994, or 1995. From these data, I was able to determine the exact sequence of courses of students who eventually enrolled in Course 4. To allow tracking data from Courses 2 and 3, I restricted the analyses to students who enrolled in Course 4 in 1994 or 1995 (N=620). As it turned out, only 20 of these students previously took Courses 1, 2, or 3. The small sample size severely limited the precision in estimating key components of the decision theory model, and as a result, prevented making firm conclusions about the effectiveness of these remedial courses. Nevertheless, the results are a useful illustration of how the model could be applied, given sufficient sample sizes.

Analyses

Questionnaire data

For each respondent, I computed a coherence indicator *CHRIND1* for the elicited grade value function (Part 2 of the questionnaire). A 'coherent' grade value function (*gvf*) is defined as one for which *gvf*(D), *gvf*(C), and *gvf*(B) are non-missing, and 0 < *gvf*(D) < *gvf*(C) < *gvf*(B) < 1. The indicator *CHRIND1* was set equal to 1, if the *gvf* was coherent, and was set equal to 0, otherwise.

I also computed a coherence indicator *CHRIND2* for the elicited course placement preferences (Part 3 of the questionnaire). *CHRIND2* was set equal to 1, if the sequence of choices in Part 3 of the questionnaire was coherent, and was set equal to 0, otherwise. (For a definition of coherence in this context, see Appendix B.)

The respondents for whom a coherent summary value function (*svf*) could be computed were those for whom *CHRIND1*=1 and *CHRIND2*=1. I therefore computed a coherence indicator for the summary value function *CHRINDS* = *CHRIND1* * *CHRIND2*.

I computed frequency distributions for the coherence indicators, and for the elicited *gvf*s and *svf*s.

Decision Theory Model

From the ACT Mathematics score and course grade data, I estimated various logistic regression models. To illustrate the expected value function results I first estimated the conditional probability of success in Course 4, given ACT scores. Success was defined in two ways: Completing Course 4 with a B or higher grade, and completing Course 4 with a C or higher grade. Under either definition of success withdrawals (W grades) and incompletes (I grades) were considered to be unsuccessful outcomes.

I estimated the conditional probability of success functions separately for the 600 students who had not taken remedial instruction before enrolling in Course 4, and for the 20 students who had. Because of the small number of students who received any kind of prior remedial instruction before enrolling in Course 4, I did not do separate analyses for particular remedial courses. By comparing the conditional probability of success curve for the students who received prior remedial instruction with the corresponding curve for students who enrolled directly in Course 4, one can make inferences about the effectiveness of the remedial instruction for students with a given ACT score.

Next, I estimated, separately for students with and without prior remedial instructions, a logistic regression model for the joint conditional distribution of course grades, given ACT Mathematics scores. I used SAS's PROC LOGISTIC (SAS Institute, 1990) to estimate an 'ordinal response model' $P[Y \leq k \mid X=x]$, where Y is the course grade, X is the ACT Mathematics score, and k corresponds to the grades A-F. I computed estimated conditional probabilities of particular grades by subtracting appropriate estimated cumulative probabilities: $P[Y=g \mid X=x] = P[Y \leq g+1 \mid X=x] - P[Y \leq g \mid X=x]$.

I then computed expected value functions. For the total group of students who enrolled in Course 4, the expected value function is:

$$E[svf] = \sum_{a=1}^{10} \sum_{x=1}^{36} svf(a) * p[a \mid x] * g(x) \tag{5}$$

where a is one of the 10 possible outcomes of the placement system; x is an ACT Mathematics score; $p[a \mid x]$ is the estimated conditional probability of outcome k, given ACT Mathematics score x; and g(x)

is the marginal probability of score x. The outcome *a* reflects both the prior remedial status as well as the final grade in Course 4 that students may have. The conditional probability function $p[a \mid x]$ can be represented by the two sets of estimated probabilities described in the preceding paragraph. The marginal probability function g(x) can be either the empirical distribution of test scores observed among all students, or it can be a smoothed or assumed distribution of test scores.

Equation (5) pertains to the decision rule currently used by Midwestern University to place students in different courses. Alternative decision rules would change the outcomes *a*, in that different groups of students would be placed in remedial courses. For example, if the decision rule were based on the ACT Mathematics test or on another placement test with cutoff score K, then $a = a(K)$ would be a function of K, and hence so would $svf=svf(K)$. We could then estimate new conditional probabilities $p[a(K) \mid x]$, and compute an expected value $E[svf(K)]$. By comparing $E[svf(K)]$ with the expected value in Equation (5), we could determine whether the alternative decision rule was better or worse than the current rule.

I have chosen here to make a more limited (and basic) comparison: How does the current rule compare to the null decision rule in which no students are placed in remedial courses (i.e., all students enroll directly in Course 4)? In this null decision rule, there are only 5 outcomes. To calculate the expected *svf*, we can simply apply the conditional probability function estimated from students who did not receive remedial instruction to *all* students who enrolled in Course 4, and use the empirical probability distribution of ACT Mathematics test scores.

Results

Value Functions

Table 3 on the following page summarizes the distribution of the coherence indicators CHRIND1, CHRIND2, and CHRINDS by respondent group. About 4/5 of the students had coherent grade value functions. About 3/4 of the students had coherent summary value functions (compared to about 2/3 of

the public university students in the Sawyer (1995) study). All 19 instructors had coherent grade value functions, and about 4/5 of them had coherent summary value functions.

Table 4, also on the following page, summarizes the medians of the elicited summary value functions by respondent group. The outcomes are denoted by letter grades followed by the symbol NR (no prior remedial instruction) or R (prior remedial instruction). Value functions for grades associated with no prior remedial instruction were also broken down by course.

Several results are immediately clear from Table 4:

* There was remarkable uniformity in the grade value functions of students in different courses.
* The median grade value functions do not resemble the threshold utility function.
* Instructors' median value functions tended to be modestly, but uniformly, higher than those of students.
* Both students and instructors would prefer that students earn a B in the standard course instead of an A if doing so meant that they would avoid taking a remedial course.

The medians of the *svf*s for the two total groups in Table 4 parallel closely the results reported by Sawyer (1995).

ERIC

Table 3.
Proportions of Respondents with Coherent Value Functions, by Group

Group	N	Grade value function	Pref. for remedial instr.	Summary value function
Students				
Course 1: Basic Algebra I	19	.68	.89	.63
Course 2: Basic Algebra II	81	.83	.90	.75
Course 3: Trigonometry	52	.88	.94	.85
Course 4: Elem. functions	109	.84	.85	.78
Total, all students	270	.83	.89	.77
Instructors				
Total, all instructors	19	1.00	.79	.79

Table 4.
Median Summary Value Functions, by Group

Group	N	Course placement outcome									
		A (NR)	B (NR)	A (R)	B (R)	C (NR)	C (R)	D (NR)	D (R)	F (NR)	F (R)
Students											
Course 1: Basic Algebra I	19	1.00	.80			.50		.10			.00
Course 2: Basic Algebra II	81	1.00	.80			.50		.10			.00
Course 3: Trigonometry	52	1.00	.80			.40		.10			.00
Course 4: Elem. functions	109	1.00	.80			.40		.10			.00
Total, all students	270	1.00	.80	.68	.56	.40	.27	.10	.07	.04	.00
Instructors											
Total, all instructors	19	1.00	.80	.70	.60	.50	.35	.20	.14	.08	.00

Logistic Regression

Table 5 on the following page shows the results of the logistic regression of success in Course 4 on ACT Mathematics score, by prior remedial instruction. Results are given for two different definitions of success: Completing Course 4 with a B or higher grade, and completing Course 4 with a C or higher grade. Under both definitions, withdrawals and incompletes are considered to be unsuccessful outcomes. The sample sizes for the logistic regression are smaller than the numbers of students enrolled in Course 4, because 3 of the students with prior remedial instruction, and 73 of the students without prior remedial instruction did not have ACT Mathematics scores.

The small sample sizes for the group with prior remedial instruction directly translates into a lack of statistical significance for the regression coefficients. Therefore, the results of this analysis and the results for the decision model can not be considered accurate enough to draw substantive conclusions about the effectiveness of Courses 1-3 in preparing students for Course 4. Nonetheless, the results do illustrate the kinds of inferences that could be made if sufficiently large samples were available.

ERIC

Table 5.
Logistic Regression Models for Success in Course 4, Given ACT Mathematics Score
(by Success Criterion and Prior Remedial Status)

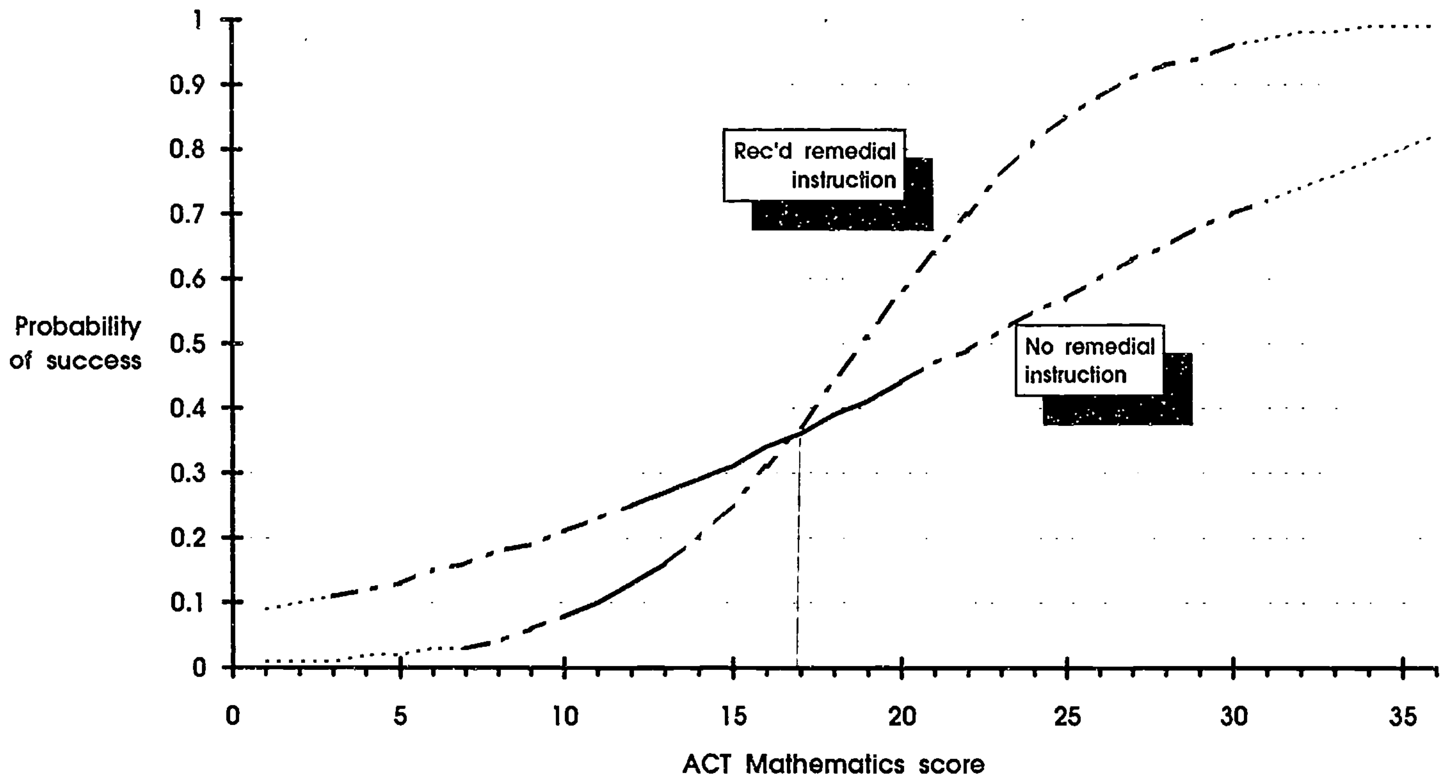
Success criterion	Prior remedial status	N	Regression coefficient	Estimate (p)
B or higher	Received prior remedial instruction	17	Intercept	-5.299 (.04)
			ACT Mathematics score	0.280 (.13)
	Did not receive prior remedial instruction	527	Intercept	-2.414 (.0001)
			ACT Mathematics score	0.109 (.0001)
C or higher	Received prior remedial instruction	18	Intercept	-2.886 (.12)
			ACT Mathematics score	0.184 (.22)
	Did not receive prior remedial instruction	527	Intercept	-1.084 (.0008)
			ACT Mathematics score	0.093 (.0001)

Figures 2 and 3 on the following pages are plots of the conditional probability functions described by Table 5. The soli.. portions of each curve correspond to the middle 50% of students; the thick dashed lines correspond to the lower and upper quartiles, and the fine dashed lines are extrapolations of the fitted curve to scores for which there were no data.

Figure 2 suggests that prior remedial instruction improves students' chances of completing Course 4 with a B or higher grade, provided their ACT Mathematics scores are greater than or equal to 17. Figure 3 (C or higher) suggests that prior remedial instruction is helpful for students with ACT Mathematics scores of 20 or higher.

Interpreted naively, Figure 2 would also suggest that students whose ACT Mathematics scores are below 17 (or below 20) are likely to do worse in Course 4 if they take remedial courses. Of course, the

Figure 2.
Probability of Success in Mathematics Course, by Prior Remedial Instruction (Success Criterion = B or Higher)

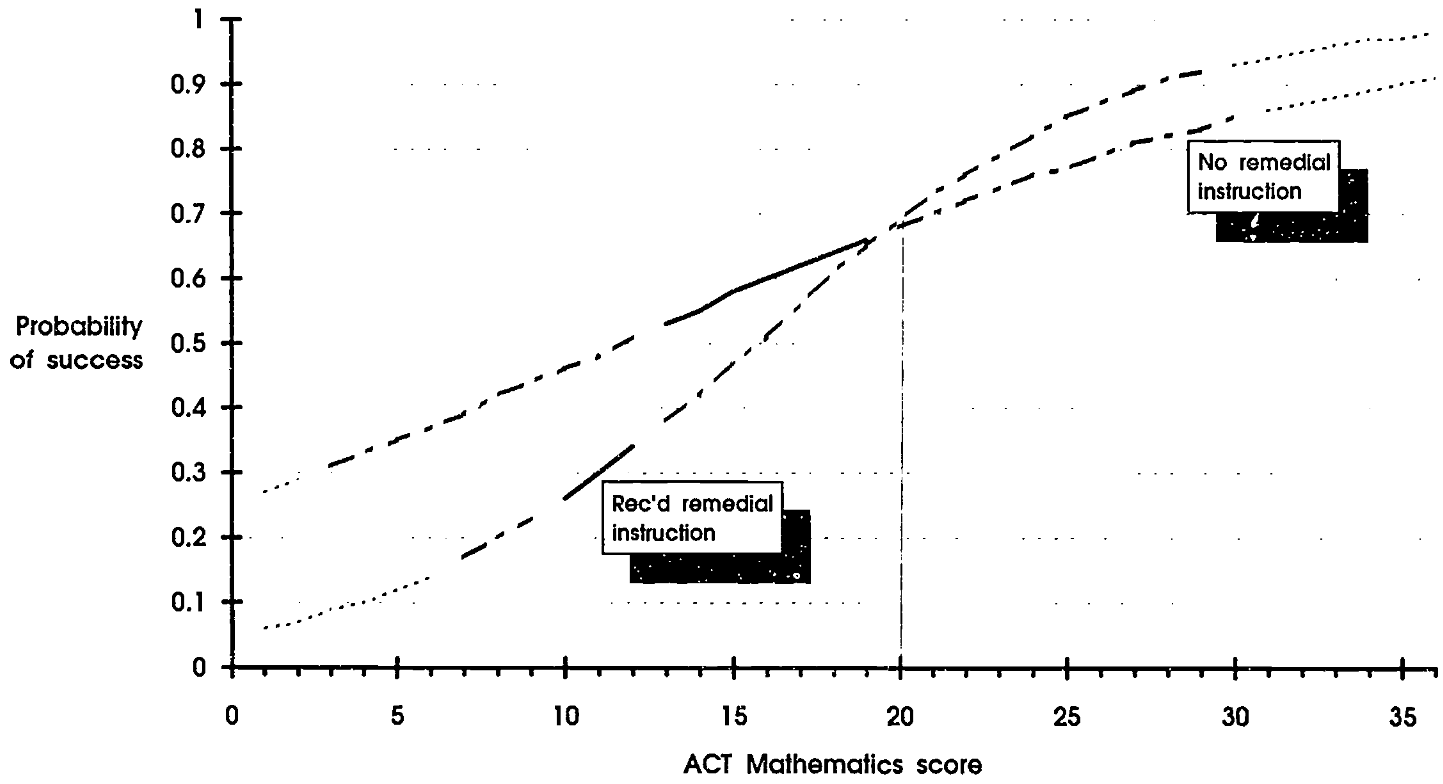


- 25 -

28

ERIC Full Text Provided by ERIC

Figure 3.
Probability of Success in Mathematics Course, by Prior Remedial Instruction (Success Criterion = C or Higher)



sampling error in estimating the probability curve for students with prior remedial instruction is so large that any differences between the two curves in this region can be attributed to chance. A curve estimated from a larger sample of students with prior remedial instruction would presumably track the other curve much more closely in the region of lower test scores.

Nevertheless, the notion that remedial instruction would benefit some students more than others is intuitively reasonable. Some students may be so poorly prepared that even taking remedial instruction does not improve their chances of succeeding in the standard course. Relatively well-prepared students may benefit from remedial instruction, and better-prepared students may benefit even more. This phenomenon is reminiscent of 'aptitude/treatment interactions' that have been the focus of much study in previous decades.

Note that the solid portion of the probability curve for students who received remedial instruction lies over the range of test scores from about 10 to 13, which is well below the 'cross-over' scores noted in the preceding discussion. If it were based on a larger sample, this result would suggest that the remedial instruction for Course 4 was largely ineffective for the students who received it.

ERIC

Expected Summary Value Functions

In the decision theory model to which the elicited *svf* applies, there are 10 outcomes, determined by prior remedial instruction and by the final grade in Course 4. Table 6 below summarizes the fitted conditional probability distributions for the course grades A-F, given ACT Mathematics scores. Separate estimates are shown by prior remedial instruction status.

Note that as in Table 5, the small sample number of students with prior remedial instruction guarantees that none of the results for that group are statistically significant. Nevertheless, I will use the results to illustrate the calculation of the expected *svf*s.

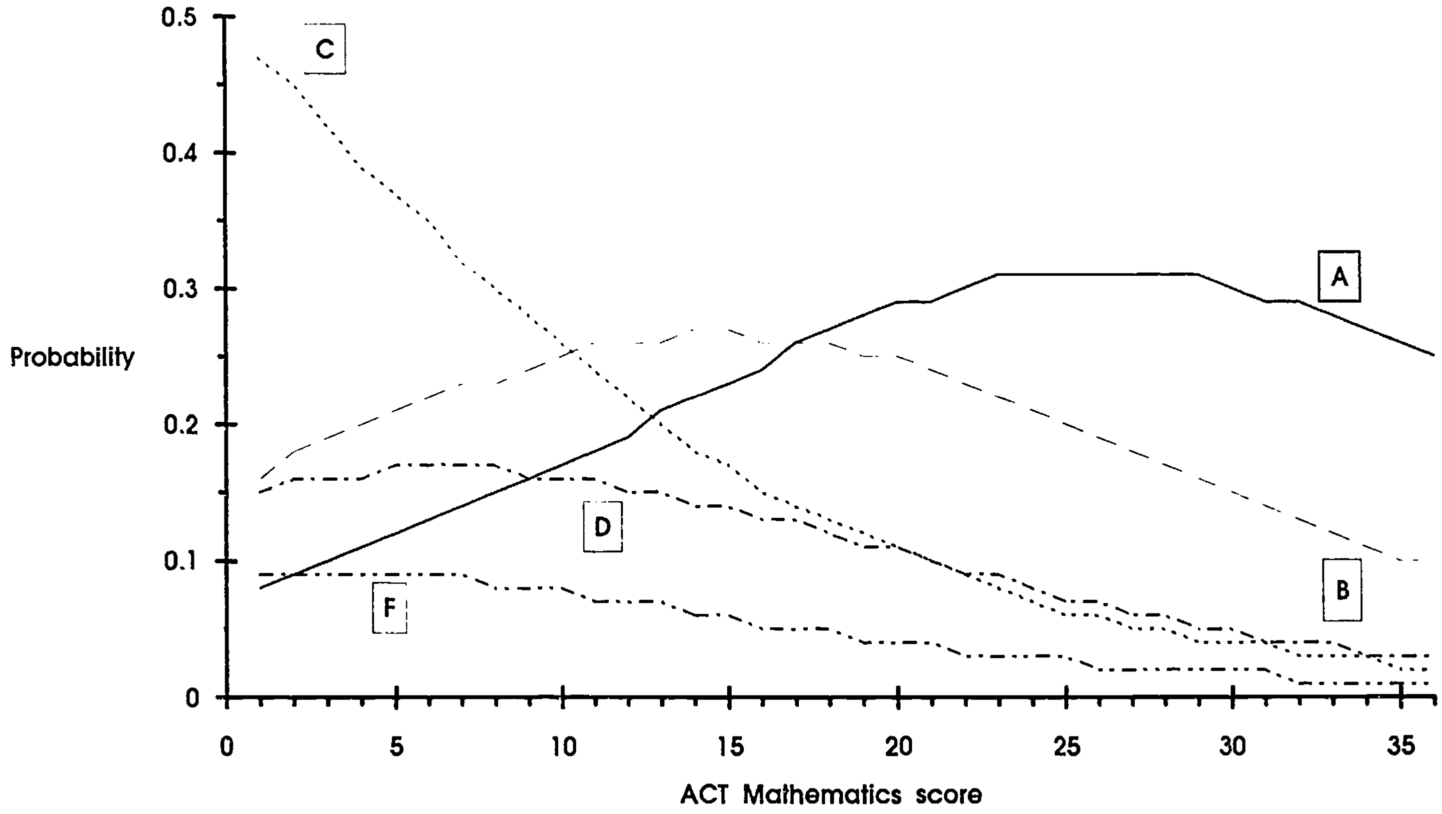
Table 6.
Estimated Conditional Probabilities of Grades in Course 4, Given ACT Mathematics Scores
(by Prior Remedial Instruction)

Group	N	$p[A \mid x]$	$p[B \mid x]$	$p[C \mid x]$	$p[D \mid x]$	$p[F \mid x]$
Prior remedial instruction	17	f(-4.6592;x)- f(-2.8980;x)	f(-2.8980;x)- f(-1.9133;x)	f(-1.9133;x)- f(-1.3937;x)	f(-1.3937;x)- f(0;x)	f(0;x)
No remedial instruction	526	g(-3.4329;x)- g(-2.1353;x)	g(-2.1353;x)- g(-1.0458;x)	g(-1.0458;x)- g(-0.3741;x)	g(-0.3714;x)- g(0;x)	g(0;x)

Notes: x = ACT Mathematics test score.
$f(t;x) = 1/[1 + \exp(-t-.1254*x)]$.
$g(t;x) = 1/[1 + \exp(-t-.1068*x)]$.
Withdrawals and incompletes were converted to F.

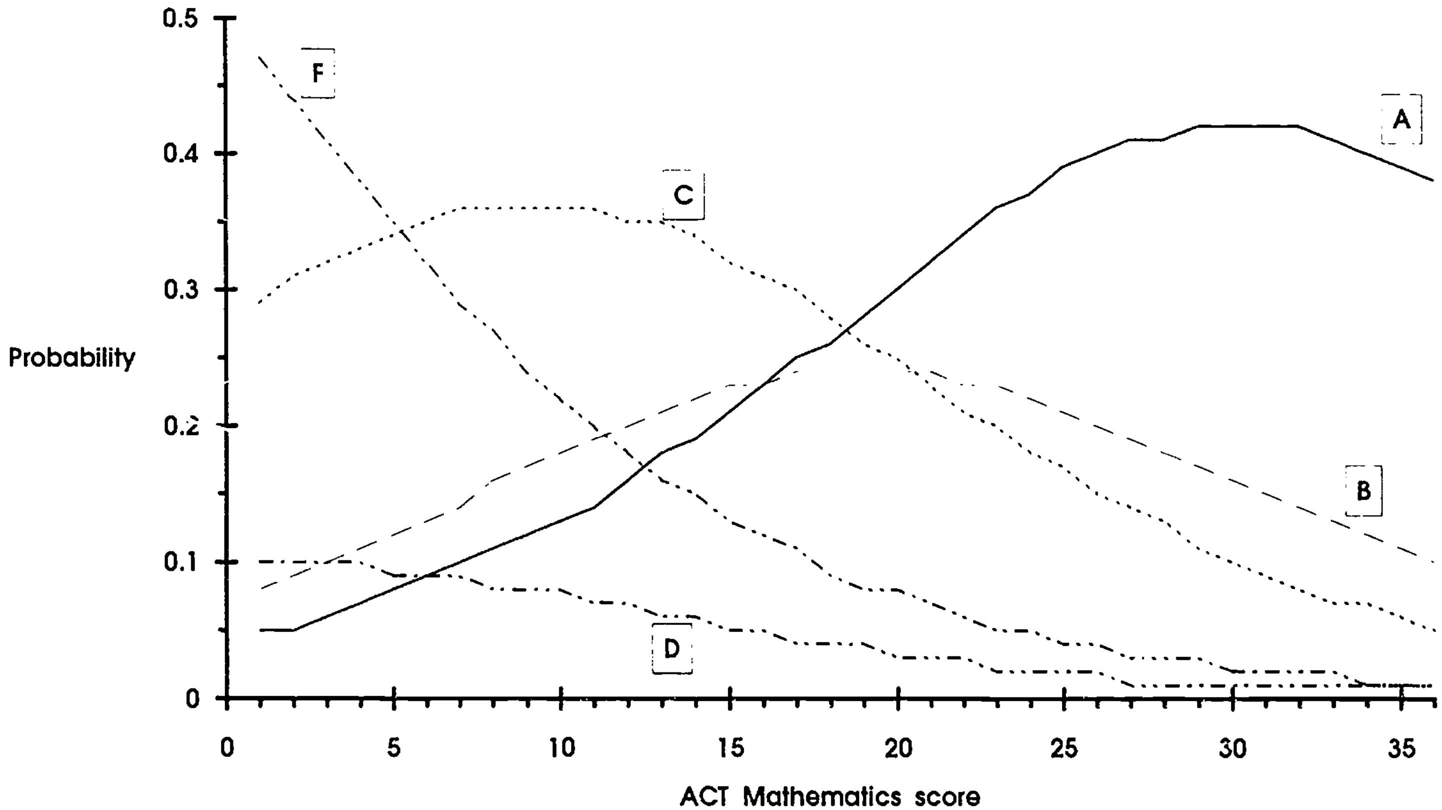
Figures 4 and 5 on the following two pages are plots of the conditional probability functions shown above. Note that the modal probabilities for particular grades are associated with different regions of the ACT Mathematics score scale.

Figure 4.
Conditional Probability of Grades in Course 4, Given ACT Mathematics Scores (Students Without Prior Remedial Instruction)



ERIC

Figure 5.
Conditional Probability of Grades in Course 4,
Given ACT Mathematics Scores
(Students With Prior Remedial Instruction)



Using the elicited *svfs* of students in Table 4, the estimated conditional probability functions reported in Table 6, and the empirical distribution of ACT Mathematics scores, I computed expected value functions associated with the following two decision rules.

* *Rule 1*: The current placement system based on high school course work and scores on the local mathematics placement test.
* *Rule 2*: A null placement system, in which all students enroll directly in Course 4.

The expected *svf* associated with Rule 1 was .504, and the expected *svf* associated with Rule 2 was .509. This result indicates that the current placement system is not achieving its goals, as expressed by the *svf* ---- in fact, it is counter-productive! This result is not surprising, given the probability of success results depicted in Figures 2 and 3, and given the penalty the *svf* assigns to each grade if a student has previously taken remedial instruction. One must keep in mind, of course, that these results are based on a very small sample of students who took remedial instruction. Therefore, the conclusion should be viewed as an illustration, rather than as a substantive conclusion about the effectiveness of course placement in mathematics at Midwestern University.

Conclusions

A college course placement system includes both an assessment component and an instructional component. The effectiveness of the system as a whole depends on both components. Statistical decision theory can be used to describe the possible outcomes of course placement systems. Building a decision theory model requires an appropriate preference function and an appropriate probability distribution for the possible outcomes.

In a study at a public midwestern university, about 3/4 of the student respondents, and about 4/5 of the instructor respondents, were able to supply coherent grade value functions. The median value functions for the instructors were modestly, but systematically higher than those for students. This result is consistent with results obtained in previous research.

Logistic regression can be used to model the conditional probability distribution of placement system outcomes, given scores on a placement test. By averaging the value function according to the conditional probability distribution and the marginal distribution of test scores, one can calculate an expected value function. By comparing expected value functions associated with different decision rules, one can make inferences about the effectiveness of an existing or of alternative placement systems.

Future research

During the next year, I plan to investigate the feasibility of eliciting course placement value functions more directly than in this study. For example, I will ask respondents to assign values to all 12 outcomes associated with taking remedial instruction and with performance in the standard course. Direct elicitation would eliminate the need for interpolation in constructing the summary value function.

I also plan to collect and analyze additional course grade data. I believe that Courses 8 and 9 have much larger enrollments than Course 4, and that they have larger pools of students with prior remedial instruction. The latter is especially important in obtaining stable estimates of the conditional probability distributions used to calculate expected value functions.

References

ACT (in press). *Technical manual for the ACT Assessment.* Iowa City, Iowa: Author.

American Council on Education (1996). *Remedial education: An undergraduate student profile.* Washington, DC: Author.

Ben-Shakhar, G., Kiderman, I., & Beller, M. (1994). *Comparing the utility of two procedures for admitting students to liberal arts: An application of decision theoretic models.* Jerusalem: National Institute for Testing and Evaluation.

The Carnegie Foundation for the Advancement of Teaching (1988). *An imperiled generation: Saving urban schools.* Princeton, NJ: Princeton University Press.

Carriuolo, N. (1994, April 13). Why developmental education is such a hot potato. *The Chronicle of Higher Education*, Sec. 2, pp. 1-2.

College Entrance Examination Board (1977). *On further examination.* New York: Author.

Colleges and universities offering remedial instruction and tutoring. (1994, April 13). *Education Week*, XIII, No. 29, p.6.

Cronbach, L. J., & Gleser, G. C. (1965). *Psychological tests and personnel decisions.* Urbana, Illinois: University of Illinois Press.

DeGroot, M. H. (1970). *Optimal statistical decisions.* New York: McGraw-Hill.

Falmagne, J.-C. (1985). *Elements of psychophysical theory.* (Oxford Psychology Series No. 6). New York: Oxford University Press.

Farquhar, P.H. (1984). Utility assessment methods. *Management Science*, 30(11), 1283-1300.

Frisbie, D. (1982). Methods of evaluating course placement systems. *Educational Evaluation and Policy Analysis*, 4 (2), 133-140.

Hancher, M. (1994, Feb. 2). The consequences of grade inflation. [Letter to the editor]. *Chronicle of Higher Education*, p. B-5.

Houston, W. M. (1993, April). *Accuracy of validity indices for course placement systems.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta.

Isaacs, G. L. & Novick, M. R. (1978). *Manual for the computer-assisted data analysis (CADA) monitor.* Iowa City, Iowa: The University of Iowa Dept. of Educational and Psychological Foundations.

Lindley, D. V. (1972). *Bayesian statistics: A review.* (Regional Conferences Series in Applied Mathematics.) Philadelphia: Society for Industrial and Applied Mathematics.

Lively, K. (1993, Feb. 24). States step up efforts to end remedial courses at 4-year colleges. *The Chronicle of Higher Education*, p. A-28.

Mac Donald, H. (1994). Downward mobility. The failure of open admissions at City University. *City Journal*, Summer 1994, pp. 10-20.

McNabb, T. (1990). *Course placement practices of American postsecondary institutions.* (ACT Research Report Series 90-10). Iowa City, Iowa: American College Testing.

Munday, L. (1976). *Declining admissions test scores.* (ACT Research Report No. 71). Iowa City, Iowa: American College Testing.

National Commission on Excellence in Education (1983). *A nation at risk: The imperative for educational reform.* (USED Stock No. 065-000-00177-2). Washington, DC: U.S. Government Printing Office.

Novick, M. R. & Lindley, D. V. (1979). Fixed-state assessment of utility functions. *Journal of the American Statistical Association, 74*, 306-311.

Petersen, N.S. & Novick, M. R. (1976). An evaluation of some models for culture-fair selection. *Journal of Educational Measurement, 13*(2), 3-29.

SAS Institute (1990). *SAS/STAT users' guide*, Fourth ed., Vol. 2. Carey, NC: Author.

Sawyer, R. (1994, April). *Eliciting Utility Functions for Validating Course Placement Systems.* Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans.

Sawyer, R. (1995, April). *Eliciting Utility Functions for Validating Course Placement Systems (II).* Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco.

Sawyer, R. (in press). Decision theory models for validating course placement tests. *Journal of Educational Measurement.*

Schiel, J., & Noble, J. (1993, April). *The effects of data truncation on estimated validity indices for course placement.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta.

Shea, C. (1994, Jan. 5). Grade Inflation's Consequences. *The Chronicle of Higher Education*, p. A-45.

Singal, D. J. (1991, November). The other crisis in American education. *The Atlantic, 268*, pp. 59-69.

van der Linden, W. (1991). Applications of decision theory to test-based decision making. In Hambleton, R. L. & Zaal, J. N. (Eds.), *Advances in educational and psychological testing.* Boston: Kluwer.

Willingham, W. (1974). *College placement and exemption.* New York: College Entrance and Examination Board.

Woods, J. E. (1985). *Status of testing practices at two-year postsecondary institutions.* Iowa City, Iowa: American College Testing and the American Association of Community and Junior Colleges.

Wright, D. A. & Cahalan, M. W. (1985, April). *Remedial/development studies in institutions of higher education policies and practices, 1984.* Paper presented at the annual conference of the American Educational Research Association, Chicago. (ERIC Document Reproduction Service No. ED 263-828).

Whitney, D. R. (1989). Educational admissions and placement. In Linn, R. L. (Ed.), *Educational measurement* (3rd ed.). (pp. 515-526). New York: ACE and Macmillan.

Yates, J. F. (1990). *Judgment and decision making.* Englewood Cliffs, NJ: Prentice-Hall.29

ERIC

Appendix A

Questionnaires

ERIC
Full Text Provided by ERIC

Students' Satisfaction With Grades and Course Placement Decisions

A research project by Richard Sawyer, ACT,
with the cooperation of Midwest University
Fall, 1995

Purpose of this Study

Part of ACT's work involves helping students decide which courses to take. I want to learn about the things you think about in making decisions about your courses.

I will ask you some questions about your academic work at Midwest University, and about your preferences for different grades and course placement decisions. This questionnaire is not a test --- there are no right or wrong answers. The questions are grouped into three parts. As soon as you finish one part, please continue on directly to the next part. The entire questionnaire should take about 5 minutes to complete.

Your instructor will distribute this questionnaire in class. Please take it home, answer the questions, and then bring it back to the next class meeting.

The information you give will be used to enhance the services ACT provides to students in the future. I sincerely appreciate your cooperation.

Part 1
Background Information

1. Please enter your Social Security Number (SSN) here: ________________________.
The Registrar's office will use your SSN to append transcript information to your responses. They will then remove your SSN from the data they return to me. I will not know your identity.

2. Please check (✓) the appropriate boxes to indicate whether you have taken, or are currently enrolled in, any of the courses in the table below. Also indicate either the grade you received in the course (if you have already taken it), or the grade you expect to receive (if you are currently enrolled in the course).

Course	**Check here if you have already taken**	**Grade you received**	**Check here if you are currently enrolled**	**Grade you expect to receive**
Course 1: Basic Algebra I				
Course 2: Basic Algebra II				
Course 3: Trigonometry				
Course 4: Elementary Functions				
Other mathematics courses (please specify):				

3. What is your major or program at Midwest University? (If you have not yet selected a major, please write "None.")

____

ERIC

4. When did you first start taking courses at Midwest University?

_______________ (month and year)

5. When did you first select your *current* major or program at Midwest University?

_______________ (month and year)

6. What is your gender?

___ Female

___ Male

7. What is your age?

_____ years

8. Which of the following statements best describes your goals about the grades you earn in courses at Midwest University? *(Check one only.)*

___ I don't mind earning a few Ds, so long as I receive credit for all my courses.

___ It is important for me to earn only As, Bs, or Cs in my courses.

___ It is important for me to earn only As or Bs in my courses.

___ It is important for me to earn all As in my courses.

Part 2
Course Grades

Students want to earn as high a grade in a course as they can. Naturally, everyone would be more satisfied with an A than with a B, or with a B than with a C, and so forth --- but what about your *relative* satisfaction? Would you, for example, feel twice as satisfied with an A as with a B?

I want to find out your relative satisfaction with grades in the courses you take. In answering the questions, please think of any course that you need to pass to satisfy the requirements of your program at Midwest University.

The line below is meant to suggest your relative satisfaction with the different letter grades. The letter grade of F is associated with 0% satisfaction, and the letter grade of A is associated with 100% satisfaction:

F **A**

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Please indicate on this line your relative satisfaction with the grades of B, C, and D by writing them above an appropriate point on the line. For example, if you would be about half as satisfied with a B as with an A, then you would write a "B" above the 50% mark.

NOTE: Your responses should reflect your relative satisfaction with particular grades in a standard course. Your responses do not have to correspond to a percent-correct grading scale (for example, where the grade A represents 90% or more correct).

Part 3
Course Placement

Let a "standard course" be a for-credit course that is required for your program. For example, some students may need to pass *Course 4: Elementary Functions* to satisfy the requirements of their program at Midwest University.

One purpose of a course placement system is to determine whether a student is ready to take a particular standard course. If a student is not ready to take the standard course, he or she can instead enroll in a "developmental course" to acquire the skills needed to succeed in the standard course. At Midwest University, for example, *Course 3: Trigonometry* would be considered a developmental course for the standard course *Course 4: Elementary Functions.*

Taking a developmental course will tend to increase a student's chances of success in the standard course. However, taking a developmental course also has disadvantages---it increases the time required to complete your program, it costs additional money, and it may not carry credit toward your degree. Therefore, the decision to take a developmental course involves a trade-off: an increased chance of eventually succeeding in the standard course, versus extra time and money.

I want to find out how *you* see these trade-offs.

The table below presents different situations in which you are asked to choose between either taking a developmental course *before* taking the standard course [Col. (1)], or *directly* enrolling in the standard course [Col. (2)]. Assume that the developmental course is 1 semester in length.

For each situation, please check (✓) either the box in Col. (1) or the box in Col. (2), according to your preference:

Situation	Which would you prefer ? Col. (1) Take the developmental course *before* taking the standard course. Then, earn this grade in the standard course:	Col. (2) Enroll directly in the standard course, and earn this grade:
1	A	B
2	A	C
3	A	D
4	A	F
5	B	C
6	B	D
7	B	F
8	C	D
9	C	F
10	D	F

*** This is the end of the questionnaire. Thank you for your help! ***

Students' Satisfaction With Grades and Course Placement Decisions

A research project by Richard Sawyer, ACT,
with the cooperation of Midwest University
Fall, 1995

Purpose of this Study

Part of ACT's work involves helping students decide which courses to take. I want to learn about the things you think about in making decisions about your courses.

I will ask you some questions about your academic work at Midwest University, and about your preferences for different grades and course placement decisions. This questionnaire is not a test --- there are no right or wrong answers. The questions are grouped into three parts. As soon as you finish one part, please continue on directly to the next part. The entire questionnaire should take about 5 minutes to complete.

Your instructor will distribute this questionnaire in class. Please take it home, answer the questions, and then bring it back to the next class meeting.

The information you give will be used to enhance the services ACT provides to students in the future. I sincerely appreciate your cooperation.

Part 1
Background Information

1. Please enter your Social Security Number (SSN) here: .
The Registrar's office will use your SSN to append transcript information to your responses. They will then remove your SSN from the data they return to me. I will not know your identity.

2. Please check (✓) the appropriate boxes to indicate whether you have taken, or are currently enrolled in, any of the courses in the table below. Also indicate either the grade you received in the course (if you have already taken it), or the grade you expect to receive (if you are currently enrolled in the course).

Course	Check here if you have already taken	Grade you received	Check here if you are currently enrolled	Grade you expect to receive
Course 1: Basic Algebra I				
Course 2: Basic Algebra II				
Course 3: Trigonometry				
Course 4: Elementary Functions				
Other mathematics courses (please specify):				

3. What is your major or program at Midwest University? (If you have not yet selected a major, please write "None.")

ERIC

4. When did you first start taking courses at Midwest University?

   ______________ (month and year)

5. When did you first select your *current* major or program at Midwest University?

   ______________ (month and year)

6. What is your gender?

   ___ Female

   ___ Male

7. What is your age?

   _____ years

8. Which of the following statements best describes your goals about the grades you earn in courses at Midwest University? *(Check one only.)*

   ___ I don't mind earning a few Ds, so long as I receive credit for all my courses.

   ___ It is important for me to earn only As, Bs, or Cs in my courses.

   ___ It is important for me to earn only As or Bs in my courses.

   ___ It is important for me to earn all As in my courses.

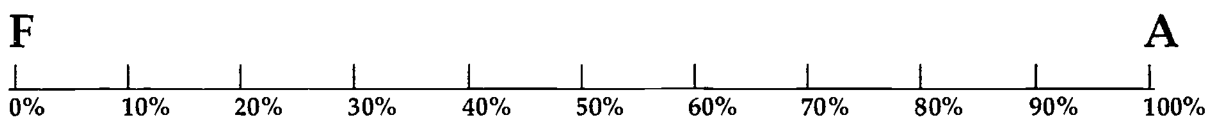
ERIC

Part 2
Course Grades

Students want to earn as high a grade in a course as they can. Naturally, everyone would be more satisfied with an A than with a B, or with a B than with a C, and so forth --- but what about your *relative* satisfaction? Would you, for example, feel twice as satisfied with an A as with a B?

I want to find out your relative satisfaction with grades in the courses you take. In answering the questions, please think of any course that you need to pass to satisfy the requirements of your program at Midwest University.

The line below is meant to suggest your relative satisfaction with the different letter grades. The letter grade of F is associated with 0% satisfaction, and the letter grade of A is associated with 100% satisfaction:



Please indicate on this line your relative satisfaction with the grades of B, C, and D by writing them above an appropriate point on the line. For example, if you would be about half as satisfied with a B as with an A, then you would write a "B" above the 50% mark.

NOTE: Your responses should reflect your relative satisfaction with particular grades in a standard course. Your responses do not have to correspond to a percent-correct grading scale (for example, where the grade A represents 90% or more correct).

Part 3
Course Placement

Let a "standard course" be a for-credit course that is required for your program. For example, some students may need to pass *Course 4: Elementary Functions* to satisfy the requirements of their program at Midwest University.

One purpose of a course placement system is to determine whether a student is ready to take a particular standard course. If a student is not ready to take the standard course, he or she can instead enroll in a "developmental course" to acquire the skills needed to succeed in the standard course. At Midwest University, for example, *Course 3: Trigonometry* would be considered a developmental course for the standard course *Course 4: Elementary Functions.*

Taking a developmental course will tend to increase a student's chances of success in the standard course. However, taking a developmental course also has disadvantages---it increases the time required to complete your program, it costs additional money, and it may not carry credit toward your degree. Therefore, the decision to take a developmental course involves a trade-off: an increased chance of eventually succeeding in the standard course, versus extra time and money.

I want to find out how *you* see these trade-offs.

ERIC

The table below presents different situations in which you are asked to choose between either taking a developmental course *before* taking the standard course [Col. (1)], or *directly* enrolling in the standard course [Col. (2)]. Assume that the developmental course is 1 semester in length.

For each situation, please check (✓) either the box in Col. (1) or the box in Col. (2), according to your preference:

Situation	Which would you prefer ?	
	Col. (1) Take the developmental course *before* taking the standard course. Then, earn this grade in the standard course:	Col. (2) Enroll directly in the standard course, and earn this grade:
1	A	B
2	A	C
3	A	D
4	A	F
5	B	C
6	B	D
7	B	F
8	C	D
9	C	F
10	D	F

***** This is the end of the questionnaire. Thank you for your help! *****

Appendix B

Constructing Summary Value Functions for Course Placement Outcomes

Appendix B

Constructing Summary Value Functions for Course Placement Outcomes

Part 4 of each questionnaire asks respondents to make 10 choices. Each choice involves either taking the remedial course before taking the standard course, and earning grade G_1 in the standard course, or else taking the standard course directly, and earning grade G_2, where $G_1 > G_2$.

The result of the choices is a sequence of Rs and Ss, where:

R= Prefer to take the remedial course before taking the standard course.
S= Prefer to take the standard course directly.

There are $2^{10} = 1024$ possible sequences of response patterns, but most of them are "incoherent," because they are inconsistent with the transitivity property of preference relations. A coherent sequence is one that satisfies the following inequalities:

a. (R,A) ≻ (R,B) ≻ (R,C) ≻ (R,D) ≻ (R,F), and
b. (S,A) ≻ (S,B) ≻ (S,C) ≻ (S,D) ≻ (S,F),

where ≻ is a respondent's preference. Then Inequality a. implies, for example, that if (S,C) ≻ (R,A), then (S,C) ≻ (R,B), because (R,A) ≻ (R,B). Moreover, Inequality b. implies that if (S,C) ≻ (R,A), then (S,B) ≻ (R,A), because (S,B) ≻ (S,C). To simplify matters, I have also assumed that the following preferences exist:

c. (R,A) ≻ (S,F)
(R,B) ≻ (S,F)
(R,C) ≻ (S,F)
(R,D) ≻ (S,F)

The inequalities in c. imply that in Choices 4, 7, 9, and 10, the respondent must always choose taking the remedial course and earning a passing grade in preference to taking the standard course directly and receiving an F. These preferences may not actually be true of students who are very willing to take risks. Making these assumptions, however, considerably reduces the number of allowable sequences. Finally, I assume that:

d. (S,F) ≻ (R,F)

Inequalities a. - d. imply that every other course placement result is preferable to (R,F) (i.e., taking the remedial course and then receiving an F in the standard course). Table 1 on p. 3 shows the 14 choice sequences that satisfy these inequalities. I computed for each respondent an indicator function CHRIND2:

CHRIND2=1, if the respondent's sequence of choices was one of those listed in Table 1; and CHRIND2=0, otherwise. The respondents for whom a coherent summary value function could be imputed were those for whom both CHRIND1=1 (where CHRIND1 is the coherence indicator for the grade value function) and CHRIND2=1. These people were identified by the summary value function coherence indicator CHRINDS=CHRIND1*CHRIND2.

Table 1.
Coherent Choice Sequences

Choice sequence	Choice number									
	1 (R,A) or (S,B)	2 (R,A) or (S,C)	3 (R,A) or (S,D)	4 (R,A) or (S,F)	5 (R,B) or (S,C)	6 (R,B) or (S,D)	7 (R,B) or (S,F)	8 (R,C) or (S,D)	9 (R,C) or (S,F)	10 (R,D) or (S,F)
1	R,A	R,A	R,A	R,A	R,B	R,B	R,B	R,C	R,C	R,D
2	R,A	R,A	R,A	R,A	R,B	R,B	R,B	S,D	R,C	R,D
3	R,A	R,A	R,A	R,A	S,C	R,B	R,B	R,C	R,C	R,D
4	R,A	R,A	R,A	R,A	S,C	R,B	R,B	S,D	R,C	R,D
5	R,A	R,A	R,A	R,A	S,C	S,D	R,B	S,D	R,C	R,D
6	S,B	R,A	R,A	R,A	R,B	R,B	R,B	R,C	R,C	R,D
7	S,B	R,A	R,A	R,A	R,B	R,B	R,B	S,D	R,C	R,D
8	S,B	R,A	R,A	R,A	S,C	R,B	R,B	R,C	R,C	R,D
9	S,B	R,A	R,A	R,A	S,C	R,B	R,B	S,D	R,C	R,D
10	S,B	R,A	R,A	R,A	S,C	S,D	R,B	S,D	R,C	R,D
11	S,B	S,C	R,A	R,A	S,C	R,B	R,B	R,C	R,C	R,D
12	S,B	S,C	R,A	R,A	S,C	R,B	R,B	S,D	R,C	R,D
13	S,B	S,C	R,A	R,A	S,C	S,D	R,B	S,D	R,C	R,D
14	S,B	S,C	S,D	R,A	S,C	S,D	R,B	S,D	R,C	R,D

Note: The shaded cells correspond to choosing to take the standard course directly.

BEST COPY AVAILABLE

If we use the customary grades A-F to measure achievement in the standard course, and if we neglect Withdrawal (W) grades, then there are 10 possible final outcomes of the placement system: X = {(R,A), . . ., (R,F), (S,A), . . ., (S,F)}, where R denotes taking the remedial course before taking the standard course, and S denotes taking the standard course directly. The set X, together with the set of possible placement test scores is the outcome space Θ.

In principle, one could elicit a value function for X with a diagram like that in Part 2 of the questionnaires. With such a diagram, however, the respondent would have to mark 8 outcomes (rather than the 3 outcomes A, B, and C) above the 0-100 scale. I believe that most respondents would have great difficulty doing this. Therefore, I elected to *impute* a summary value function *svf* for X, using the grade value function *gvf* elicited in Part 2 of the questionnaires as a reference. Now, there are many ways one could impute a summary value function; I chose the simplest method I could think of. Specifically, the imputed value function *svf* has the following properties:

a. *svf*(S,G) = *gvf*(G), for G = A, B, C, D
b. *svf*(R,F) = 0.
c. *svf*(S,F) = .25**gvf*(D)
d. For G = A, B, C, D, the values of *svf*(R,G) are interpolated between appropriate values of 1=*gvf*(A), *gvf*(B), *gvf*(C), *gvf*(D), and 0.

Equation a. says that the summary value function associated with taking the standard course directly and earning a particular grade G is equal to the grade value function *gvf* elicited in Part 2 of the questionnaires. Equation b. says that the worst possible result is to take the remedial course, then receive an F in the standard course. Equation c. says that taking the standard course directly, and receiving an F is slightly better than receiving an F in the standard course after taking the remedial course; I have arbitrarily assigned the value .25**gvf*(D) to this result. Property d. says that the outcomes associated with first taking the remedial course are to be assigned values according to the respondent's 10 choices in Part 4 of the questionnaires. Provided that the respondent's sequence of choices is one of the 14 coherent sequences listed in Table 1, it is possible to interpolate between values of *gvf* in a consistent way. Each of the 14 coherent choice sequences defines a separate imputed summary value function *svf*. The resulting values of the imputed summary value functions *svf* are shown in Table 2 on the following page.

Table 2.
Imputed Summary Value Functions for Course Placement Outcomes

Choice sequence	First course (Grade in standard course) Std. (A)	Rem. (A)	Std. (B)	Rem. (B)	Std. (C)	Rem. (C)	Std. (D)	Rem. (D)	Std. (F)	Rem. (F)
1	1	.50 + .50*gvf(B)	gvf(B)	.50*gvf(B) + .50*gvf(C)	gvf(C)	.50*gvf(C) + .50*gvf(D)	gvf(D)	.50*gvf(D)	.25*gvf(D)	0
2	1	.50 + .50*gvf(B)	gvf(B)	.50*gvf(B) + .50*gvf(C)	gvf(C)	.75*gvf(D)	gvf(D)	.50*gvf(D)	.25*gvf(D)	0
3	1	.50 + .50*gvf(B)	gvf(B)	.75*gvf(C) + .25*gvf(D)	gvf(C)	.50*gvf(C) + .50*gvf(D)	gvf(D)	.50*gvf(D)	.25*gvf(D)	0
4	1	.50 + .50*gvf(B)	gvf(B)	.75*gvf(C) + .25*gvf(D)	gvf(C)	.75*gvf(D)	gvf(D)	.50*gvf(D)	.25*gvf(D)	0
5	1	.50 + .50*gvf(B)	gvf(B)	.90*gvf(D)	gvf(C)	.75*gvf(D)	gvf(D)	.50*gvf(D)	.25*gvf(D)	0
6	1	.75*gvf(B) + .25*gvf(C)	gvf(B)	.50*gvf(B) + .50*gvf(C)	gvf(C)	.50*gvf(C) + .50*gvf(D)	gvf(D)	.50*gvf(D)	.25*gvf(D)	0
7	1	.75*gvf(B) + .25*gvf(C)	gvf(B)	.50*gvf(B) + .50*gvf(C)	gvf(C)	.75*gvf(D)	gvf(D)	.50*gvf(D)	.25*gvf(D)	0
8	1	.75*gvf(B) + .25*gvf(C)	gvf(B)	.75*gvf(C) + .25*gvf(D)	gvf(C)	.50*gvf(C) + .50*gvf(D)	gvf(D)	.50*gvf(D)	.25*gvf(D)	0
9	1	.75*gvf(B) + .25*gvf(C)	gvf(B)	.75*gvf(C) + .25*gvf(D)	gvf(C)	.75*gvf(D)	gvf(D)	.50*gvf(D)	.25*gvf(D)	0
10	1	.75*gvf(B) + .25*gvf(C)	gvf(B)	.90*gvf(D)	gvf(C)	.75*gvf(D)	gvf(D)	.50*gvf(D)	.25*gvf(D)	0
11	1	.75*gvf(C) + .25*gvf(D)	gvf(B)	.50*gvf(C) + .50*gvf(D)	gvf(C)	.25*gvf(C) + .75*gvf(D)	gvf(D)	.50*gvf(D)	.25*gvf(D)	0
12	1	.75*gvf(C) + .25*gvf(D)	gvf(B)	.50*gvf(C) + .50*gvf(D)	gvf(C)	.75*gvf(D)	gvf(D)	.50*gvf(D)	.25*gvf(D)	0
13	1	.75*gvf(C) + .25*gvf(D)	gvf(B)	.90*gvf(D)	gvf(C)	.75*gvf(D)	gvf(D)	.50*gvf(D)	.25*gvf(D)	0
14	1	.90*gvf(D)	gvf(B)	.75*gvf(D)	gvf(C)	.60*gvf(D)	gvf(D)	.50*gvf(D)	.25*gvf(D)	0

Note: Choice Sequence No. 1 corresponds to always choosing to take the remedial course before taking the standard course. The shaded cells for Choice Sequences Nos. 2 - 14 show the modifications in the summary value function that are associated with choosing to take the standard course directly.

BEST COPY AVAILABLE