DOCUMENT RESUME

ED 414 973 JC 980 045

AUTHOR Rasor, Richard A.
TITLE Evaluating a College Success Course.
PUB DATE 1998-01-00
NOTE 23p.; Funded by Statewide Student Success Course Initiative. Prepared in cooperation with the Research and Planning Group for California Community Colleges.
PUB TYPE Reports - Evaluative (142)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Basic Skills; Community Colleges; *Course Evaluation; *Developmental Studies Programs; Educational Research; *Evaluation Methods; Evaluation Needs; Evaluation Research; Experimental Groups; *High Risk Students; *Program Effectiveness; Research Projects; *Study Skills; Two Year Colleges

ABSTRACT

This paper addresses the methodological issues encountered when attempting to evaluate a college course's effectiveness by determining whether students who enroll in it perform at higher levels than students who did not enroll. The paper identifies resource people who may be helpful in the evaluation, and offers guidelines for the research question, which defines the purpose of the project, and the null or research hypothesis. The necessity of operational definitions is also emphasized, with examples and definitions of key words. Suggested factors to consider in planning the research project include cost, ethics, time, choice of students for the study, availability of students, commitment, faculty, the experimenter effect, student retention during the study, outcome measures and when to examine them, treatment rigor, and necessary sample size. The four research models provided for evaluation of the college success course are: (1) the experiment with randomly assigned partially matched students; (2) the experiment with randomly assigned unmatched students; (3) the post-facto design; and (4) the time series design. Selection of a particular model depends on the collective purpose of the evaluation. The paper concludes with a suggested outline for the ensuing research report.(YKH)

Evaluating a College Success Course

by

Richard A. Rasor, Ed.D.
American River College

in cooperation with:

The Research and Planning Group
for
California Community Colleges

Funded by:

The Statewide Student Success Course Initiative

Program Director: Dr. Ed Stupka
Sacramento City College
(916) 558-2231

January 1998

Table of Contents

Evaluating a College Success Course

Abstract

This paper addresses the methodological issues encountered when attempting to evaluate the effectiveness of a specific course designed to improve overall student academic performance in community colleges. Also included are four research models which are designed to answer the question of whether students who enroll in such a course subsequently perform at higher levels than students who do not take the "success" course. Topics include: obtaining assistance from resource people on campus, defining the research purpose, stating the research hypothesis, operational definitions, costs, ethics, which students to use, which faculty to use, time lines, the experimenter effect, student survival while "in treatment," and outcome measures. The research models are: the experiment with randomly assigned partially matched students, the experiment with randomly assigned unmatched students, the post-facto design, and the time series design.

Evaluating a College Success Course[1]

There are many factors to consider and several ways to evaluate the impact of a course that is specifically designed to increase academic success for community college students. Some evaluation designs are relatively simple and easy to accomplish depending upon one's resources. Other designs are more complex depending upon how tightly one wishes to isolate the effects of the success course from all the other factors that can impact a student's academic life and are likely to also play a role in academic performance. In this paper, I shall try to keep in mind that some people are fortunate by having access to an existing database, are knowledgeable about research design and statistics, and have full cooperation from other faculty, staff, and administration. In other instances, a person may have little access, few research skills, and will have to function in relative isolation. Whatever the case may be for you, try to get some assistance (assuming you need it) *at the very beginning of the planning phase of your research.* Having such help at the beginning of your project can help protect you from making many mistakes commonly seen. Here are some people likely to be found at your campus or district office that can offer valuable assistance in helping you get started:

Resource People

1. **The college or district researcher**. This person will most likely maintain a computer database of student files (demographics, courses, final grades). If you can avoid rummaging through student folders during your research, do it. With a database, you may be able to deal with an entire population of students rather than with a small sample. The researcher will also be knowledgeable about design and probably have access to statistical software for analyzing your results.

2. **An instructor of research design and/or statistics**. Check the psychology department. Most likely these instructors of psychology have had some academic training in research methods and statistics. They might also have some good reference books on these subjects. Often the math department offers a statistics course and may also have suitable software for statistical analysis. Check with those instructors too.

3. **Data processing personnel / instructors**. If there is no researcher to process your data, you will definitely want someone from data processing to help you with setting up files and possibly with data analysis. Data processing staff may not have training in research methodology and statistics, but they would certainly know how to identify a particular group of students (a cohort) and track their progress.

4. **A reference librarian**. Whether on your campus or at a nearby university, reference librarians can provide invaluable assistance with locating materials and with helping conduct literature searches (e.g., professional journals). Before you begin your research, it is wise to find out what other people have done with the same general topic. Try an ERIC search.

5. **A person with good internet skills.** While this person could also be a reference librarian, often there are others who have developed this skill to a high degree. Browsing the web has potential for answering any question you may put to people. For example, using the internet, one researcher asked

---

[1] Appreciation is extended to James E. Barr of American River College for his review and assistance in the preparation of this manuscript.

if anyone had experience with a particular sampling technique. Many helpful responses were received!

6. **An instructor or counselor known for exceptional achievements.** Because you are given the responsibility of evaluating a specific course, it is very helpful to have, as a resource person, an instructor or counselor who is not directly connected to the program, department, or course in question. Sometimes an "outsider" can provide insights often overlooked by those who have had considerable experience with this type of course or program.

7. **A person representing the administration at your college**. This person will be able to provide you with the administration's perspective on all the matters related to enrollment or registration, as well as present your views to the administration. This person may also have the power to help with "red tape."

8. **Others**. By now you are probably thinking that the resource people previously mentioned would make a good steering committee for your project. Correct! If you do want such a committee, also include a student or two, and a representative from classified personnel. It is a good idea to involve a representative of the academic senate or at least keep the senate informed of your plans and progress.

Once your resource people are in place, it is time to start detailed planning. At your institution, I assume that you already have a college success course, or about to start offering one, and need to evaluate its impact. To have you think about this project like a researcher does, imagine that the college success course is like a "medical treatment," given to some individuals but not offered or denied to others. Maybe even think of the success course as a prescription pill versus a false pill, a placebo. How would medical authorities know if the real pill truly affected a disease it was being used to treat? Some diseases simply go away without any treatment. And it is also known that just taking a placebo can be beneficial too, provided of course, that people think it to be a real treatment and not just a placebo. Similarly, the success course is definitely a treatment, given to some individuals and not to others. Now if students who take the success course perform academically better overall than students who did not take this course, is it because of the course itself or could it be due to one of many other factors having little to do with the course? The whole thrust of scientific research is to seek truth about cause and effect. It's not an easy task.

Defining the Purpose of the Project: The Research Question

One of the first things you need to do in planning your project is to define the research purpose. You may do this by forming a question. Here are some typical research questions as related to evaluating the impact of a success course:

- Does the success course result in higher subsequent student grades and persistence?
- To what extent does the success course contribute to higher academic performance?
- Has student success and persistence increased since implementing the success course?
- Does the success course cause higher academic performance in students?

Notice that these example purpose statements or research questions have considerable similarity. The essence of them is simply, is this specific course causing a change in student performance? There are many variations on how to word a specific research question. The final wording depends upon what people are comfortable with in terms of describing what the research is about. In this instance, the primary focus is upon evaluating academic outcomes as a function of a success course. Now you might say, "No need to bother with this, I know what I what to do." So why is the research question desirable? Having developed a specific question helps to clarify what the research is intended to answer. Such a question gives one focus when deciding upon all the variations or options you will face in preparing a research study. It is easy to get lost in methodological detail and analysis. With a clear research question constantly in mind, this is less likely to happen.

Stating the Null Hypothesis or the Research Hypothesis

Having formed a research question, the next thing that many people do is form a hypothesis. This is done in advance of data collection (or data analysis), and is a logical extension of any research question. Many people are apt to tell you that a hypothesis is simply a guess about how things will turn out. Not quite true! A hypothesis is taken seriously by researchers. Like the research question, a hypothesis helps keep the focus upon what the research is all about. As such, it may even be more specific in terminology than a research question. *A hypothesis is a statement about the outcome, but it is not simply a wild guess.* Much thought goes into it, and a researcher usually has already reviewed the professional literature (e.g., professional journals or ERIC) to determine what others have already found.

The null hypothesis (null means zero in German) is quite simply, there is no difference between the groups being compared (or that the difference is zero). For example, the null hypothesis might be stated as, *there is no difference in second term grade-point average between freshman students who complete a college success course and freshman students who do not complete such a course.* With the null hypothesis, there is no directionality implied because the researcher would reject the hypothesis if there was *any* type of a substantial difference between the groups, i.e., a higher *or* lower grade-point average. When a researcher fails to state a hypothesis but still uses statistical analysis, the null is automatically assumed.

One may also form what is called a research hypothesis which has a sense of directionality to it. For example, a research hypothesis might be, *freshman students who complete a college success course have a higher second term grade-point average than freshman students who do not complete such a course.* This hypothesis has directionality because it stated that one group comes out higher on grade-point average. Researchers use such directional hypotheses when there is interest in only one direction of the outcome. In the example, if students who completed the college success course had an equal or a substantially lower grade-point average than students not completing the course, there would be no support for the research hypothesis as it was stated. Finally, any hypothesis must be falsifiable, in other words, capable of being shown to be incorrect. A hypothesis which cannot be shown to be false is worthless.

Defining Your Terms: Operational Definitions

Ever hear people at your college use the word "retention?" Certainly. Now, were they really talking about how many students are left in a course at the end of the term? Were they referring to how many students are still attending the college at the end of the term? Perhaps they were thinking of re-enrollment for the subsequent term, that is, persistence. You can see the problem. Words in the English language have much elasticity of meaning. Yet in research, we must be very exact about what is meant in terms of how we measure it. Operational definitions, then, specify the operations needed to produce the terminology. For example, "The temperature is hot." The operations needed to reach this conclusion is to look at a thermometer and specify something like "The temperature is at least 100 degrees." Older students may be defined as any student age of 30 years or more. *Notice that whatever operational definitions are agreed upon, they are clearly measurable.* They leave no doubt about meaning although people may still quarrel with your selection of a particular operational definition. Here are some terms and operational definitions that may be useful:

Treatment: The event or situation that has been administered to one group of students and not to others. The treatment is also referred to as the independent variable. In some research designs, an experimental group receives the treatment while the control group does not.

Level of Treatment: The degree or amount of treatment that is administered to any individual or group of individuals, e.g., doseage. Number of units enrolled may be thought of as a level, e.g., <12 versus 12+. Being in a short-term versus long-term treatment program may also be thought of as a level.

Treatment Course: The success course that is to be evaluated regardless of its formal name. Calling it the treatment course helps to avoid confusion with certain academic outcomes also called "success" or "successful."

Student Attributes Any student demographic such as age, gender, or ethnicity may be an attribute. Student assessment score, unit load, course history, student opinion, or other background information are also considered student attributes. You may also call these simply student characteristics. Use the same categories that are usually reported at your college.

Successful A student is successful in a course with an end-of-term grade of *A, B, C,* or *CR* (credit).

Unsuccessful A student is unsuccessful in a course with an end-of-term grade of *D, F, NC* (no credit), *I* (incomplete) or *WT* (withdrew from the course after census date and the notation appears on the transcript). Grades of *IP* (in progress), as used with open ended courses, are not used for purposes of determining success.

Student Success Rate The simple count of successful grade notations divided by all grade notations (excluding *IP*) multiplied by 100 and rounded to the nearest whole number. Units associated with specific courses are not used in calculating the student success rate.

Course Success Rate The simple count of all successful grade notations given in the course divided by all grade notations given in the course (excluding *IP*) multiplied by 100 and rounded to the nearest whole number.

College Success Rate The total count of successful grade notations divided by all grade notations (excluding *IP*) multiplied by 100 and rounded to the nearest whole number. Units associated with specific courses are not used in calculating the college success rate.

Course Retention	A student is retained in a course if the transcript shows an end-of-term grade of *A, B, C, CR, D, I,* or *IP*. (Note: Retention is highly correlated with success. Researchers are recommending that success be used in place of retention).
College Retention	A student is retained in college if, at the end of term, the transcript shows at least one final grade of *A, B, C, CR, D, I,* or *IP*. (Note: Retention is highly correlated with success. Researchers are recommending that success be used in place of retention).
Grade-Point Average (GPA)	Using only grades of *A* through *F*, a student's grade points are determined by multiplying the unit value of each course by the value associated with the grade earned, e.g., A = 4, and then summed for all courses. Total grade points are then divided by total units attempted to yield GPA.
Persistence	A student is counted as having persisted to the subsequent term if there is any final grade notation on the transcript (even an *F* or *WT*) for the beginning and the subsequent term.
Persistence Rate	Using a specific count of students as a base, the persistence rate is the percentage of those students who persisted to the subsequent term.

Factors to Consider in Planning the Research Project

<u>Costs</u>. While you are probably familiar with the major costs of a research project, the smaller items sometimes go unnoticed. Pencils, paper, typing, duplicating, stamps, gasoline, and food are just some items frequently overlooked. Try to anticipate the costs for all the little things. Allow some room in your budget for unexpected contingencies.

<u>Ethics</u>. Any researcher is bound to encounter a troublesome issue which touches upon ethics. The primary consideration for you to keep in mind is, *how will my research project affect its participants?* Here are some very important matters to decide upon in advance: (1) Any student (or other administrative persons) likely to be involved in the study must be given enough pertinent information about all aspects of the research to make an informed decision as to whether to participate. This does <u>not</u> mean that you have to state your research question or hypothesis for all to know. What it does mean is that you must provide enough information about what any participant is likely to encounter during the execution of the research process. Withholding pertinent information or giving false or misleading information which would likely affect a decision to participate is a violation of research ethics. (2) No participant shall be subject to physical, mental, or emotional harm as part of the research process. (3) It is <u>not</u> a violation of ethics to deny a treatment (e.g., the treatment course) for some while giving it to others. Any new treatment should be evaluated against some standard. Often that standard is a control group which is denied treatment, at least temporarily. This last point raises an important issue in an educational setting. If the treatment is a specific course, can you legally and ethically deny registration for some students while granting it for others? If not, then you may have to consider that only certain selected students or certain sections of the course are part of the treatment process. Another way to deal with this issue is to be very clear with the pool of likely participants right from the start that *some students will be selected and some will not due to enrollment restraints. Furthermore, those students not selected may be given the option of enrolling at a later date.* Should the selected students know if they are "in treatment?" In a perfect experimental situation, the answer is "No." However, in educational situations, it is nearly impossible to

disguise enrollment in a course. The best thing to do is withhold information about the research question and hypothesis from all students.

Time. If you are starting a new research venture, allow a substantial block of time for planning (e.g., a semester). Administrative staff and faculty who are involved will have to eventually agree upon every aspect of the research - not just on general principles but upon detail as well. If you have the financial means and allowable time frames, consider running the research project first as a "pilot test." This would mean that a second semester is devoted to discovering all the little problems that are bound to surface the first time something new is actually put into practice. The third semester is the "real test." Having the benefit of a full academic year for planning and pilot testing, the third semester is ideal to run the official study.

Which Students to Use in the Study. If the treatment course is already in existence, or if it is in the proposal stage, my recommendation is the same: *For researching the effectiveness of the course, use new freshman students without prior college units*. This helps to control for differences in background which may create problems with interpreting the results of your study. Also, freshman students are the "life blood" of community colleges, and collectively, they also have the lowest academic success rates. As such, they are the group most academically at-risk. At American River College, the first-term success rate for new freshmen without previous college experience is about 60% and for all other students 70%. The freshmen also have a persistence rate of 60%, meaning that 40% do not re-enroll the subsequent semester. It seems logical to design a "success" course mostly for these students. Of course, this does not preclude further restricted enrollment in the treatment course such as all new freshman students with an assessment test score qualifying them for placement into non-transfer level courses.

Availability of Students. If you use new freshman students for the evaluation of the treatment course, then you will need contact with them before they are allowed to register. The ideal procedure would be to give these prospective students some advance information about the course including the fact (if it is) that enrollment will be limited. If you think it is needed, you may also offer a small incentive for volunteering to enroll in the course such as individualized help with registration. *From the list of all volunteers, you will eventually select only half of them for enrollment in the treatment course*. The rest will not be allowed to enroll in the treatment course or they may enroll at a later date (after the study is over). Variations on this are explained in the research models to be discussed shortly.

Commitment. Have college administration commit, in advance, to the entire research protocol. For example, what if only a handful of freshman students enroll in one section of the treatment course? Will it be cancelled? Will regular students be allowed to register for it so that enrollment will be higher? These potential problems must be thought through before the study begins. Ideally, you proceed with whatever naturally happens. Also have a contingency plan should you or anyone else connected with the study drop out. Who will take over? Who will take the place of an instructor who has been awarded a sabbatical leave? You may not be able to think of every contingency, but giving some thought to back-up plans could save you from distress at a later date. Finally, once the research is initiated, you should follow through to the end, for better or for worse. Don't change things in midstream unless it is critical.

Faculty. Selecting the faculty who will be teaching the college success course during its evaluation must be done carefully. Remember that the primary purpose of your project is to determine if the treatment course makes a difference in terms of subsequent student success. Its purpose is not to determine which faculty member does the best job with the course. In short, this research project is *not* the time to encourage inter-faculty creativity. It is best if every faculty member does about the same thing in class. This would reduce the amount of variation within the course sections. You might ask, why not use just one instructor then? That would help with controlling unwanted variation, but it would also limit the degree of being able to generalize the findings from your study. You want to create a balance between tight controls and being able to reflect the real academic world as it will exist after the study.

The Experimenter Effect. You have probably heard of a double-blind study. It means that neither research participants nor the administrative staff who have contact with the participants know who is in the experimental group (treatment) and who is in the control condition (placebo). While this is ideal, it is difficult to accomplish in an educational setting because instructors of the treatment course will certainly know that their course is "the treatment." Furthermore, such instructors may unconsciously favor such students by grading them higher than students in a control group. It is also common knowledge that not all instructors of the same course have the same grading standards. A student grade of "C" from one instructor may translate to a "B" with another. Detailed planning and agreeing upon course content and exit standards is vital if there is to be more than one instructor involved. To minimize the possibility of an instructor favoring one group of students over another, give serious consideration to having all student evaluations done by an outside authority who is not concerned with the research outcome and is kept "blind" as to whether a student is in treatment or in some type of control situation. Another alternative is to have all treatment course instructors use the same objectively scored tests with agreed upon grading criteria.

Student Survival While in Treatment. Thus far I have urged that selected faculty tighten up their grading so that it is more uniform and to be aware of their effect upon producing a desired outcome (the experimenter effect). There remains another problem that has been encountered before, attrition. What if the treatment course is so difficult, or so "boring," or perceived as not being worthwhile by the enrolled students? There is bound to be some attrition, but what if it is excessive? A medical analogy (but a false one) will serve to explain: Assume that there is a marvelous surgical operation which has been shown to extend life. The only problem is that 50% of all patients die on the operating table! Similarly, a college success course may be a great thing to extend academic life - but only for some. If treatment course attrition ("death on the operating table") is high, then perhaps the treatment is only serving to create a "Darwinian survival of the fittest" and has little to do with causing greater academic success. About all you can do is make the course attractive and interesting by using good faculty, be uniform in course content, and set realistic goals for successful completion. Students who drop out or fail the success course haven't really received "treatment." Collectively, they have to be conceptualized as a different group, not a treatment group and not a control group.

Outcome Measures and When to Examine Them. There are many possible outcome measures (commonly called dependent variables) that are possible to examine in this study. However, if there are too many, the results are confusing and sometimes even contradictory. The research question and hypothesis should define the outcome measure(s). My recommendation is to examine for both groups of students, the overall success rate, non-cumulative GPA, and persistence rate. The student benefits attributed to being in the treatment course, even for a short time, may occur during the concurrent semester. However, the better period to examine outcomes, is the subsequent term following completion of the treatment course. For this reason, I recommend two check points for success rate and GPA (not cumulative GPA), the end of term 1 and the end of term 2 with the treatment being given during term 1. Remember that a student who drops out of the treatment course or even fails it, is a special case and should not be considered as having received treatment or as having experienced a control condition. Individual persistence is at least one grade notation of record for both terms (see operational definitions).

Treatment Rigor. If you are going to expend considerable energy doing an evaluation of this course, then pack the treatment with plenty of events, materials, and information that is likely to make a big difference with students. This doesn't mean the course is made needlessly difficult. Rather, the content will be highly valuable for students. For example, do you suppose that a one-unit course is likely to have the same impact as a three-unit one covering roughly the same material? One of the reasons for not finding a difference between experimental and control groups is failure to create a powerful treatment. The other is sample size.

Necessary Sample Size. If in the entire population of community college students the likely effect of a success course is relatively small, you will need a large sample to detect any difference between the treatment and control conditions. Jaccard and Becker (1997) have provided tables of sample sizes depending upon the likely effect in the population. The only way you would know about the effect in the population is to have several years of data or have an idea of what other colleges have found by means of a literature search. To give you some idea of sizes, the range of suggested sample sizes when comparing two groups that have been randomly assigned to either a treatment or a control condition, is from 30 students in each condition, (i.e., 30 successfully finishing the treatment course and 30 controls) to nearly 400 in each condition depending upon the magnitude of the effect in the population. As a compromise, I suggest you try to get 200 students in each condition. This may require more than one semester. If the data are already available from the college database, you may be able to handle an entire population rather than a small sample.

The Research Models

There are four methods or models to help answer a particular research question and they vary in the degree of control over unwanted variables. Such undesirable variables tend to create a situation where there is less of a definitive answer to the research question. All models have advantages and disadvantages. You need to pick the one which best serves the collective purpose. A good reference book on research design is Shaughnessy and Zechmeister (1990). The models are listed in approximate order from most control over unwanted variables (and most complex) to least control.

Model 1. The Experiment With Randomly Assigned Partially Matched Students

The fundamental use of this experimental design is to determine if a college success course that is taken by some students but not others, has any effect upon concurrent and subsequent academic performance. As described, the experiment attempts to show that the college success course causes a particular outcome. This particular model is very powerful but requires some additional work with the matching.

Features:

1. You will need background information on all the students who volunteer to be part of the study before enrollment is allowed. In the title to this model there is reference to partially matched students. The matching is done on some variable that is likely to affect the outcome. An assessment test score immediately comes to mind provided that this measure has been previously shown to be related to general college success. If the variable to be matched upon has little or no relationship with overall college success, there would be no basis for matching. However, if this variable does have a meaningful relationship to student outcomes, then by matching you exercise control by having equivalent groups on the matching variable. In the research world (not a natural educational environment), students would enroll in identical courses except for the treatment course, and in the subsequent semester also enroll in identical courses. This process would create nearly equivalent groups (experimental and control) with the only substantial difference being the treatment course. Thus, any difference in subsequent student performance could be attributed to that course. However, in an educational setting, we cannot control the courses that students are currently or subsequently enrolled. This means that there are many variables likely to influence student outcomes besides the treatment course. That is why the design refers to *partial matching*. Remember, though, that control over any important variable is better than no control.

2. You will need to form two groups from the student volunteers, treatment and control. The treatment group will be enrolled in the college success course while the control group will not. Students in the control group are not told that they are in a control situation, only that the selection was done in a fair and impartial manner. Those students not selected may be told that they can take the same course next semester (perhaps even guaranteed a seat). This helps alleviate what some staff may perceive as an ethics or legal problem. Notice that the control condition is a natural one in that such non-selected students are simply allowed to do what hundreds of other freshman students have done over the years prior to implementing this course. Some staff may feel that the treatment course is necessary for all students and put forth resistance to forming a control group. While their intentions are in the best interests of students, they presume that the effectiveness of the treatment course has been established when it has not been empirically demonstrated.

3. Using assessment scores as an example, the matching process is done in this way with only the volunteers: Scores are rank ordered from highest value to lowest value. They are then paired, that is, the highest two scores are paired, then the next two, etc. If one score is extreme relative to the others, you may be better off to discard that particular person from the study. Now you have a distribution of paired assessment scores. Next, flip a coin (or do some other random process), where only one score (and therefore, one student) will be assigned to the experimental group. The non-picked score will be placed into the control group. When you have finished this process, there will be two groups with equivalent assessment scores. In this way, you will control for initial differences in skill level that could influence the outcome of your study. Remember that you must match students on some variable that is important to the study. If you wish to match on more than one variable, then the pool of students from which to draw

must be larger as it is more difficult to find such matches. For example, it would be rare to find pairs of students who not only have nearly the same assessment score but also the same relative high school GPA.

4. A serious drawback to any research design is attrition. If many students drop out of treatment (or out of the control condition), the end result could be relatively small matched groups. If a student in the control condition drops out of school, that is a natural occurring event that is considered an outcome measure for this study. In any event, the goal is to have a minimum of 30 students in each group and, if they were in the treatment course, finish it successfully.

5. Suggested student outcomes to analyze for both groups are GPA, success rate, and persistence to next term. These evaluations may be done at two check points: At the conclusion of the term of the experiment and at the conclusion of the subsequent term. The GPA mean for each group as well as the mean success rate for each of the two groups may be compared using a *t* test for independent data. (A correlated *t* test is not recommended because matching is not complete in all respects and one does not want natural attrition to force equal reduction in the size of the other group). Persistence is a matter of either a person did or did not re-enroll grouped by treatment and control conditions. This will result in a 2 x 2 frequency table which can be analyzed using a chi square test independence. Correlated chi square is not recommended. These statistics are described in Jaccard and Becker (1997). Also refer to the beginning of this paper for a list of resource people should you want help with statistical analysis.

Model 2. The Experiment With Randomly Assigned Unmatched Students

Should you find Model 1 impractical, then Model 2 attempts to accomplish the same thing without having to do matching. Once again, the fundamental use of this experimental design is to determine if a college success course (the treatment), that is taken by some students but not by others, has any effect upon subsequent academic success and persistence. As described, this experiment attempts to show that the treatment causes a particular outcome. This model is powerful and it eliminates the matching process associated with Model 1.

Features:

1. You will need a pool of students who volunteer to be in the study. From these volunteers you will form two groups (treatment and control). Assign each volunteer a number with no duplicates. Using a table of random numbers or any other random procedure, select half of the volunteers. Now you will have equal or nearly equal sized groups. For example, if there were 140 volunteers, you would randomly pick 70 leaving two groups of 70 each. Now label the first half that were picked as "heads" and the other half as "tails." Now flip a coin. If it comes up heads, the first group will be considered as the treatment group, that is, allowed to register for the treatment course. If the coin comes up tails, the second group will be enrolled in the course. By double randomizing (picking half and then picking an entire group) there is no intentional bias likely to affect the outcome of the study. Complete randomization is intended to create roughly equivalent groups on important variables thereby replacing deliberate matching as was done in Model 1. It does not guarantee that the two groups are equivalent - only that there is no intentional bias.

2. Once selection has been completed, you may want to compare the averages (assessment scores or high school GPA) for the two groups to see if they are nearly the same. If they are not, you have a problem! You are not supposed to trim or adjust the selected samples, nor are you supposed to go through another random selection process. Such manipulation of the chosen sample only feeds the critic with ammunition that you adjusted things to support your hypothesis. Probably the best thing to do is simply note in your report that there was an initial difference between the groups. However, there are some complex statistical techniques that may be used to compensate for unequal groups and these are mentioned below.

3. Suggested student outcomes to analyze for both groups are GPA, success rate, and persistence to next term. These evaluations may be done at two check points: At the conclusion of the term of the experiment and at the conclusion of the subsequent term. The GPA mean for each group as well as the mean success rate for each of the two groups may be compared using a t test for independent data. Persistence is a matter of either a person did or did not re-enroll and would be grouped by treatment and control conditions. This will result in a 2 x 2 frequency table which can be analyzed using a chi square test of independence. These statistics are described in Jaccard and Becker (1997). Also refer to the beginning of this paper for a list of resource people should you want help with statistical analysis.

4. If you have a large initial difference between the two groups with respect to assessment means or high school GPA, then you cannot say that the groups were equivalent at the start of the experiment. Two procedures that may be used by a statistician when there is a non-equivalent control group, are analysis of covariance and multiple regression analysis. For either one, the statistics are best done by computer software. Analysis of covariance has the effect of adjusting the means of the outcome measures as per the initial measures (e.g., assessment scores). While this approach is not vigorously embraced by statisticians for equating groups, it is sometimes used by researchers for just this purpose. There are many

assumptions underlying the use of analysis of covariance. Multiple regression analysis does not control for anything but allows you to know how much of the variance in student outcome measures is explained by knowledge of assessment scores (or any other predictor variable). Assessment scores are entered first into the equation followed by knowledge of treatment received (treatment or control). This allows you to determine the amount of variance in student outcome as a function of treatment once assessment scores are known. Another approach is to enter assessment scores second, that is, following knowledge of treatment received. This would provide you with information on the amount of variance in student outcome that is explained by assessment scores once treatment is known. As you might well assume, these are complex matters and must be handled only by persons who are competent with these particular statistical approaches. Two good reference books on these analyses and other statistical techniques are Hinkle, Wiersma, and Jurs (1998) and Ferguson and Takane (1989).

Model 3. The Post-Facto (After-the-Fact) Design

Many times it is not possible or even desirable to randomly assign people to a treatment in order to determine if the intervention makes a difference. For example, we cannot expose people to a dangerous disease in order that we may have a pool of volunteers for random assignment to either an experimental treatment or control condition. Similarly, we cannot force some people to lose their jobs and become poor so that research can determine the effect upon psychological health. Then in situations when random assignment is not feasible, we must rely upon natural occurring outcomes, i.e., people who already have the dangerous disease or who are already out of work and poor. Thus, in post-facto research, experimental grouping is done by some characteristic or quality that these people bring to the study beforehand such as gender, age, or educational background. Hence, post-facto literally means after the fact. Other terms for the same design are *post hoc* or *ex post facto*.

Features:

1. This design can be entirely proactive in that you may compare the outcomes from students who are about to take the treatment course against those who will not. The study can also be retrospective because necessary data are already available. You then compare historical records of those who took the treatment course to those showing no record of having had the course *during the same time period*. Whether proactive or retroactive, these features apply; the "experimental group" are those students who successfully completed the treatment course while the "controls" are those students who, during the same time period, did not enroll in the course.

2. If you have access to a database which houses all records for those students who did and did not enroll in the treatment course, you may examine demographics, assessment test performances, high school GPA, loads, etc. Comparing the two groups across these variables will tell you how they differ in other respects.

3. When done retrospectively, you may also "sift" through your college's computer database and try to create some *post facto* matching of the groups provided that it is done in a impartial manner and you do not use the outcome data that will be analyzed, i.e., GPA, success, or persistence. For example, suppose you believed that gender is an important variable in your study. Further assume that 86 female and 75 male students completed the treatment course during a given semester. Then instead of comparing the treatment group against everyone else who did not take the course during the same semester, you could randomly select from your total "control" group, a sample the same size as the treatment group but which would consist of 86 females and 75 males. To do this type of after-the-fact selecting, requires that you pick only an important variable and then randomly select the necessary quota from the available students who were not in the treatment course. Also be aware that trying to create some equivalency between the groups invites criticism regarding "adjusting the data to produce a desired effect." You should show the comparison data before and after "sifting."

4. No matter how sophisticated your comparisons may be in a *post facto* study, the results must be explained as a correlated finding. The reason for this is that students self-selected their treatment (by choosing to enroll in the course) and were not randomly assigned. There is no way of knowing, then, if self-selection created a unique group and that such students may have been subsequently successful irrespective of a treatment course. An example may help to clarify this concept. I have been involved for a number of years with a program that provides tutors for specific courses. Students may or may not avail themselves of these services. If they do, they must commit to a regimen of meeting with a tutor for at

least three hours a week. In each class section that offered this choice, some students selected the tutoring while others did not. The results were dramatic. Students who received tutoring had an overall success rate of 87% in the course to 57% for those who did not enter into the tutoring arrangement but had the same instructor. Substantially higher success rates for tutored students were found across all courses for many years. Assessment test score means were the same for both groups. Can it be assumed, then, that tutoring made the difference? No. Tempting though it may be to assume this, the results are only correlated, that is, *students who received tutoring tended to have higher success rates.* This does not rule out cause and effect - simply that in any *post facto* study one cannot automatically rule in a causal explanation.

5. Suggested student outcomes to analyze for both groups are GPA, success rate, and persistence to next term. These evaluations may be done at two check points: At the conclusion of the term of the experiment and at the conclusion of the subsequent term. The GPA mean for each group as well as the mean success rate for each of the two groups may be compared using a t test for independent data. Persistence is either did or did not re-enroll and would be grouped by "treatment" and "control" conditions. This will result in a 2 x 2 frequency table which can be analyzed using a chi square test of independence. These statistics are described in Jaccard and Becker (1997). Also refer to the beginning of this paper for a list of resource people should you want help with statistical analysis.

Model 4. Time Series Design (Then and Now)

When a program has already been implemented and affects a large body of students, its effectiveness can only be evaluated by examining outcomes before and after the implementation of the program. For example, if a college success course has been required of all new freshman students for some time, and there is no interest in suspending this requirement or in creating any type of concurrent control group, then one is left with comparing the academic outcomes of such students before the course was in the curriculum (or required of all) and after its implementation. It is called a time series design because you examine two periods of time, before treatment was available and after treatment was available. Of course, different students are compared as are different time periods.

Features:

1. You will need suitable historical records from the college database. First, select a particular type of student (e.g., new freshmen without prior units) during academic terms that preceded implementation of the treatment course. Second, determine the GPA and success rate for these groups during the first and second term of enrollment. Also determine the persistence rate to the second term of enrollment. Third, keep compiling data from different semesters that preceded the treatment course until you have established a stable base rate for all the student outcome measures, e.g., two academic years.

2. Having established how your student groups performed academically before there was a treatment course, now repeat the process of combining several terms and use the same type of students who successfully completed the treatment course. Hence, you will have established two base rates, "then" and "now." Try to capture data covering the same number of semesters for both time frames.

3. Assume that the base rate for student success is 50% before there was a treatment course and that it is 65% following the implementation of the course. Can you assume that the 15% gain is caused by the treatment? No. Once again, it is a correlated finding which does not rule out cause and effect but does not automatically rule it in. Different time frames and different students were used.

4. Suggested student outcomes are combined for several semesters and include for both time periods: GPA, success rate, and persistence to next term. These evaluations are done at the conclusion of the terms preceding the treatment course and at the conclusion of the terms after the implementation of the course. The GPA mean for each period as well as the mean success rate for each of the two periods may be compared using a *t* test for independent data. Persistence is either did or did not re-enroll and would be grouped by "Then" and "Now" conditions. This will result in a 2 x 2 frequency table which can be analyzed using a chi square test of independence. These statistics are described in Jaccard and Becker (1997). Also refer to the beginning of this paper for a list of resource people should you want help with statistical analysis.

5. This design can be substantially improved by using, as a comparison group, students from another similar college that has never implemented this particular treatment course. For example, assume that freshmen from this comparison community college had a success rate of 55% before the course was implemented at *your* college. Further assume that during the period following implementation of your treatment course, this comparison (or control) college had a freshmen success rate of 57%. Thus, their success rates over these two time periods without a treatment course showed a gain of 2%. Using the example data from #3, the gain at your college was 15%. You are now in a position to suggest that 13% of the 15% gain is likely due to the treatment, because at a similar college, their gain was only 2%

without such a course. While such results are still a correlated finding, the example conclusion is greatly strengthened by having comparison data during the same time frames.

6. An alternative procedure for presenting the data is as follows: Instead of compiling data from several terms to form "then" and "now" base rates, simply record the performance measure, (e.g., success rate) for each of four terms preceding the treatment as well as each of the four terms following implementation of the treatment course. Plot these in the form of a graph. On the vertical axis of the graph, scale various success rates. On the bottom axis, scale the terms (semesters). Within the graph, plot the obtained success rates and connect the dots with a straight line. You may also plot the success rates for any comparison college within the same graph. Finally, draw a vertical line in the graph that corresponds to when the treatment course was implemented at your college. Such a graph may help communicate the visual impact of such a course.

Writing the Report

If your research project is grant supported, there is a specified manner in which the funding agency wants you to write both a proposal and the results of the study. Simply follow their guidelines. But if the project is not externally funded, you may find the following outline helpful:

Title Page and Abstract. Your title should be brief but contain the necessary language to communicate the content with the main variables. For example, "Academic Performance of Students Who Enrolled in a College Success Course." The Abstract, which follows the title page, is a short overview of the entire project. Between 100 - 200 words in length, it covers the research question, a brief discussion of methodology, and a summary of the main findings.

Introduction. This section follows the Abstract and any page of acknowledgments. Give the reader a brief history of what results have already been found with courses like this. Here is where you review the professional literature that has been published by other researchers. Cite their last name and date of report in the body of this section. Finish this section with the research question and your hypothesis.

Methodology. There are three subsections within this general topic. They are Research Participants, Materials, and Procedure.

> **Research Participants**. In this subsection you describe the students who participated in the study. Include gender, median age, summary assessment data (if appropriate), ethnicity, and other pertinent information. Provide enough detail so that another researcher who is interested in replicating your study will be able to select the same type of students.
>
> **Materials**. Describe any assessment test that was used, major handouts to students, other written materials used in solicitation for volunteers, and instruments for evaluating student performance in the treatment course. You may include full copies of documents such as these in an appendix.
>
> **Procedure**. Describe within this section, everything that was done in the study. Start with early events such as the plan for organizing the treatment course and end with later events. Include how students were selected, assigned to treatment, and how they were evaluated. This is the subsection which would enable another researcher to know everything about how you did the study - including what went wrong, if anything.

Results. The result section is usually kept short and to the point. Statistical results are presented along with decisions relative to a hypothesis. For example, was there support for a hypothesis that suggested students who took the treatment course have a greater subsequent success rate than students who did not take the course? Regardless of outcome, you need to reflect upon such a hypothesis. Either you found support or did not find support for it.

Discussion. In a discussion section, you may elaborate on such things as the design, interesting findings, and suggest changes for the next study. In short, this section allows the reader to see more of "you" as related to your research. Sometimes, if results are very short, you can combine the Results section with the Discussion section.

References. List the authors in alphabetical order. Include date, title of the research, publisher, or where a copy can be found.

Appendix. If you believe that certain documents should be shown in their entirety, then photocopies may be placed in an appendix.

This ends the discussion of designs and other matters used to evaluate your college success course. The entire process, while somewhat complex and laborious, is very worthwhile. A person who has reached a conclusion about such a course that is based upon empirical research evidence is in a far greater advocacy position than one who is operating from a hunch.

References

Ferguson, G. A. & Takane, Y. (1989). Statistical analysis in psychology and education. (6th edit.). New York: McGraw-Hill Publishing Company.

Hinkle, D. E. , Wiersma, W., & Jurs, S. G. (1998). Applied statistics for the behavioral sciences. (4th edit.). Boston: Houghton Mifflin Company.

Jaccard, J. & Becker, M. A. (1997). Statistics for the behavioral sciences. (3rd edit.). Pacific Grove, CA.: Brooks/Cole Publishing Company.

Shaughnessy, J. J. & Zechmeister, E. B. (1990). Research methods in psychology. (2nd edit.). New York: McGraw-Hill Publishing Company.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

JC980045

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: EVALUATING A COLLEGE SUCCESS COURSE

Author(s): RICHARD A. RASOR

~~Corporate~~ Source: American River College

Publication Date: 1/98

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following two options and sign at the bottom of the page.

Level 1	Level 2
[✓] Check here **For Level 1 Release:** Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical) ***and*** paper copy.	[ ] Check here **For Level 2 Release:** Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical), but ***not*** in paper copy.
The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2 documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Sign here→ please

Signature: Richard A. Rasor

Printed Name/Position/Title: RICHARD A. RASOR RESEARCH COORDINATOR

Organization/Address: AMERICAN RIVER College 4700 College Oak Dr SACRAMENTO, CA 95841

Telephone: (916) 484-8166

FAX: (916) 484-8519

E-Mail Address:

Date: 1/21/98

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:	Jonathan Kelly ERIC Clearinghouse for Community Colleges 3051 Moore Hall Box 951521 Los Angeles, CA 90095-1521

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to: