PLATO®

Guidelines for Evaluating PLATO® Programs

TRO Technical Paper #2

Rob Foshay, Ph.D.
Vice President - QA,
Standards & Training

A learning resource from

TRO Learning, Inc.

4660 West 77th Street, Edina, MN 55435

(612) 832-1000

http:www.tro.com

foshay@tro.com

October, 1992

Table of Contents

Introduction

The PLATO® system gives you a powerful tool to improve learning. The system can be used in many ways, for many different purposes. It is sound educational practice to evaluate the effectiveness of any instructional program, including PLATO implementations. But it's important to use an evaluation design that is appropriate to the way in which you are using PLATO, and to your needs for decision-making. When evaluations are inconclusive, it's often the case that an inappropriate evaluation design was used, and the desired result *could not* have emerged from the evaluation.

This paper will outline six basic evaluation designs. For each, we'll describe the type of PLATO application for which it is appropriate, the questions the evaluation design can answer, the requirements for successful evaluation, the procedures for data gathering, and the procedures for data analysis. The six evaluation designs are:

- Mastery-Based Program Effectiveness
- Non-Mastery-Based Program Effectiveness
- Affective Outcomes
- Case Studies / Qualitative Assessments
- Program Comparison

- Cost Comparisons

At the end are a summary table of evaluation procedures and some guidelines for writing the evaluation report. The Appendices include a sample data analysis and sample user questionnaires.

To decide what kind of evaluation to do, you must first know who will use the evaluation results and what decisions they need to make. You can then combine and adapt these basic designs to meet your program's needs. Look at the "Questions" and "Requirements" sections for each evaluation design to identify the one(s) closest to your situation. Please feel free to contact TRO to discuss your evaluation plan.

You may want to assemble an "evaluation task force" to help you plan and execute the evaluation. Members should include:

- Key decision makers who will use the conclusions from the evaluation
- Instructors involved in the study
- The evaluator
- TRO support personnel
- Your PLATO lab manager

The group should meet at least once during evaluation planning, periodically during the evaluation, and again when the preliminary findings are ready for discussion.

Mastery-Based Program Effectiveness

In most mastery-based programs, PLATO is the primary means of initial instruction, and learners are managed so that learners begin a learning activity only when they are ready for it, and they continue to study the activity until they have fully mastered it. Thus, in contrast to a conventional instructor-led classroom, achievement is fixed but instructional time varies. As an example, mastery-based instruction is typical of GED programs.

Questions

Use this design when you want to find out:

- How well are the learners progressing due to PLATO?
- What is the longest, shortest, and average time to mastery?
- What proportion of learners master the skills taught at the required level?
- What proportion of learners do not complete or drop out?

If PLATO is the primary instructional mode in a mastery-based program, learning outcomes can be attributed mostly to PLATO.

In addition to these questions, people sometimes wish to know about long-term retention of what's learned. Unfortunately, research on long-term retention indicates that it begins as a product of initial level of mastery and the way in which topics are taught, but that by far the biggest single factor influencing long-term retention is whether the learner has an opportunity to use what's been learned. Consequently, there is no simple way to evaluate long-term retention due to PLATO or any other type of instruction.

Requirements

The program must be a true mastery model program with these features:

- Learners match the PLATO target population profile:
  - Young adult or adult
  - 3d grade minimum reading level
  - No major learning disabilities
  - General intelligence within normal limits
- Learners have virtually unlimited access to PLATO.
- Learners are placed at appropriate starting points using the assessment tests, module mastery tests, *FASTRACK* or an external norm-referenced test to which PLATO has been successfully aligned (for example, CTB-McGraw-Hill's TABE).
- Each learner receives an individual prescription which is reflected in his or her routing activity. Prescriptions may be manual or may be generated automatically via *FASTRACK* or *PLATO 2000 S.T.A.R.*
- Learners work individually at their own pace, undistracted.
- 1-1 peer tutoring is encouraged, at the instructor's discretion.
- Learners can use the module pretests to "place out" of a module and go on.
- Learners must pass each module's mastery test before going on.
- If a non-PLATO norm-referenced test such as TABE is used upon completion, it must:
  - be a parallel form of the same test used for placement (to allow reporting of gains)
  - be successfully aligned to the PLATO curriculum
  - be administered within three days of program completion
- All non-PLATO instructional activities (such as tutoring, small-group activities, off-line reading or assignments) are synchronized with the corresponding PLATO module, so that the learner participates only when he or she is studying that topic. (This requirement usually precludes large-group instruction.)
- All non-PLATO instructional activities are aligned for content with the appropriate PLATO modules. This includes using identical or complementary objectives, and compatible explanations of skills (particularly in math).

- (Recommended) Tutors are available regularly during PLATO study time to answer questions and monitor individual progress by using the available reports.
- (Recommended) Learners receive individual progress summaries periodically.

For valid statistics, complete data must be obtained for at least 25 learners. Data may be combined across times or sites, if it is acceptable to state conclusions in terms of the program as a whole but not individual times or sites.

Procedures for Data Gathering

WHEN BEGINNING the evaluation:

1. Give each learner a unique name or logon ID and use it on all data collection, on-line and off-line.

2. Administer the placement test (using a norm-referenced test such as TABE) and save the results in raw score form if available. *Note:* if you use FASTRACK for placement, you still must use the mastery tests or an external test for pretest scores.

3. Monitor implementation of the program to verify that all of the conditions of a mastery program are in fact being met. Problems commonly occur in areas such as:

- Equipment or installation difficulties/delays
- Grading policy
- Instructor training
- Instructor scheduling, availability and participation
- Alignments
- Prescriptions
- Saving data in correct form
- Learner scheduling or availability/transfers
- Learner profile

- Disincentives to succeed
- Timely post-testing

If you observe problems in any of these areas, note them for later use in interpreting the outcomes.

AS EACH LEARNER COMPLETES the program:

4. Print out CMI reports on each learner's mastery status, number of tries, and time on task for each course.

5. Administer the post-test, if any, within three days of completion.

Procedures for Data Analysis

(See the Appendix for an example of this analysis.)

IF YOU ARE USING AN EXTERNAL TEST:

1. Record beginning and ending raw test scores separately for each learner in each curriculum (reading, language arts, math).

2. If you are using a norm-referenced test, convert each score into percentiles, stanines or grade level equivalents.

3. Compute a gain score by subtracting each pretest from each post-test score (after conversion, if any).

4. Sort and then regroup the data into not more than 10 ranges, so that the average range has at least 5 learners in it.

5. Plot three paired-bar graphs for math, reading and language arts. Each graph should compare number of learners showing pre- and post-test scores in each range.

6. If required, compute the average (mean) gain.

7. If required, perform a *t*-test of statistical significance to compare the average gain against the null hypothesis of no gain.

IF YOU ARE USING PLATO'S INTERNAL COMPETENCY-BASED MASTERY TESTS ONLY:

1. Record mastery status for each learner in each course.

2. Compute an average percentage of learners who master across each curriculum.

3. Plot three bar graphs showing percent who mastered in each course in each curriculum. Superimpose a horizontal line showing the curriculum-wide average.

4. Using the completion data, add another bar next to each of the course mastery bars you drew in Step 3. The bars you add should show percent in the corresponding achievement bar who completed the course.

5. Prepare another set of bar graphs showing time on task for each course, ranging the data so that an average of at least 5 learners are in each bar.

6. Separate learners who mastered from those who did not. Then repeat step 5 to compare the two groups by preparing a paired-bar graph comparing time on task for each group. Range the data so there is an average of 5 learners per bar.

If you prefer, you can draw line graphs instead which compare time on task with gain or mastery status.

7. On the graph you drew in Step 5, mark the longest, shortest and average time to mastery.

Non-Mastery-Based Program Effectiveness

In a non-mastery program, instructor-led classes typically are the primary means of initial instruction. Although PLATO could serve in this role, it is more common for PLATO to be used in a complementary or supplementary way. Learners are managed so that instructional time (pacing) is fixed and achievement varies.

Note: intermixture of self-paced PLATO and fixed-paced classroom instruction is not recommended and will produce an uninterpretable evaluation outcome.

Questions

Use this design when you want to find out:

- How well are the learners progressing in the program as a whole?
- What is the longest, shortest, and average time on PLATO?
- What proportion of learners master skills at the required level?
- What proportion of learners do not complete or drop out?

Note that since this kind of program mixes PLATO and other instructional modes, the effects of PLATO cannot be easily separated and identified using this evaluation design. Therefore, any conclusions you draw must be about the program as a whole, not about PLATO in particular. If you want to draw quantitative conclusions specifically about PLATO vs. other instructional methods, see the "Program Comparison" design, below. You may also wish to use the "Case Study/ Qualitative" design, below, to gain a better understanding of how PLATO interacts with the non-PLATO parts of the curriculum.

In addition to these questions, people sometimes wish to know about long-term retention of what's learned. Unfortunately, research on long-term

retention indicates that it begins as a product of initial level of mastery and the way in which topics are taught, but that by far the biggest single factor influencing long-term retention is whether the learner uses what's been learned in the interim. Consequently, there is no simple way to evaluate long-term retention due to PLATO or any other type of instruction.

Requirements

The program can be any instructor-led (fixed paced) program with these features:

- Learners match the PLATO target population profile:
  - Young adult or adult
  - 3d grade minimum reading level
  - No major learning disabilities
  - General intelligence within normal limits
- Learners have regular scheduled access to PLATO (at least 3-4 hours/week is recommended).
- (Recommended in most situations) all learners use the same routing activity.
- Learners work individually and undistracted, but all work on the same PLATO module at once, in synchrony with the non-PLATO (classroom) activities. Learners should be allowed whatever time they need to finish the PLATO module. If they finish early, they can leave or be assigned an unrelated "enrichment" activity such as a PLATO life skills lesson, a computer game, or an off-line activity not related to the basic curriculum.
- (Recommended) 1-1 peer tutoring is encouraged at the instructor's discretion.
- (Recommended) Learners can use the module pre/post tests to "place out" of a module and go on.
- Learners must pass each module's mastery test before going on.
- If you use a non-PLATO norm-referenced test such as TABE upon completion, it must be:
  - a parallel form of the same test used for placement (to allow reporting of gains)
  - successfully aligned to the curriculum, including PLATO

- administered within three days of program completion

- All non-PLATO instructional activities are synchronized with the corresponding PLATO module, so that the learner participates only when he or she is studying that topic.
- All non-PLATO instructional activities are aligned for content with the appropriate PLATO modules. This includes using identical or complementary objectives, and compatible explanations of skills (particularly in math).
- (Recommended) Tutors are available regularly during PLATO study time to answer questions and monitor individual progress by using the available reports.
- (Recommended) Learners receive individual progress summaries periodically.

For valid statistics, complete data must be obtained for at least 25 learners. Data may be combined across times or sites, if it is acceptable to state conclusions in terms of the program as a whole but not individual times or sites.

Procedures for Data Gathering

WHEN BEGINNING the evaluation:

1. Give each learner a unique name or logon ID and use it on *all* data collection, on-line and off-line.

2. Administer the pre-test (using a norm-referenced test such as TABE) and save the results in raw score form if available.

Note: You may prefer to use a criterion-referenced test, such as a state test of skill levels, instead. However, this will prevent you from showing grade-level gains.

3. Monitor implementation of the program to verify that all of the conditions for successful program implementation are in fact being met. Problems commonly occur in areas such as:

- Equipment or installation difficulties/delays
- Grading policy

- Instructor training
- Instructor scheduling, availability and participation
- Alignments, especially of PLATO and non-PLATO activities
- Saving data in correct form
- Learner scheduling or availability/transfers
- Learner profile
- Disincentives to succeed

If you observe problems in any of these areas, note them for later use in interpreting the outcomes.

AS THE CLASS COMPLETES the program:

4. Print out CMI reports on each learner's mastery/completion status, number of tries, and time on task for each course.

5. List the learners who completed and who dropped out or did not complete the course as a whole.

6. Administer the post-test, if any, within three days of completion.

Procedures for Data Analysis

(See the Appendix for an example of this analysis.)

A. IF YOU ARE USING AN EXTERNAL NORM-REFERENCED TEST:

1. Record beginning and ending raw test scores separately for each learner in each curriculum (reading, language arts, math).

2. Convert each score into percentiles, stanines or grade level equivalents.

3. Compute a gain score by subtracting each pretest from each post-test converted score.

4. Regroup the data into not more than 10 ranges, so that the average range has at least 5 learners in it.

5. Plot three paired-bar graphs for math, reading and language arts. Each graph should compare number of learners showing pre- and post-test scores in each range.

6. If required, compute the average gain.

7. If required, perform a *t*-test of statistical significance to compare the average gain against the null hypothesis of no gain.

Be sure to phrase all conclusions so they apply to the program as a whole, and NOT to individual components such as PLATO.

B. IF YOU ARE USING COMPETENCY-BASED MASTERY TESTS ONLY, SUCH AS PLATO'S INTERNAL MASTERY TESTS OR MOST STATE COMPETENCY TESTS:

1. Record mastery status for each learner in each course.

2. Compute an average mastery percentage across each curriculum.

3. Plot three bar graphs showing percent mastered in each course in each curriculum. Superimpose a horizontal line showing the curriculum-wide average.

Be sure to phrase all conclusions so they apply to the program as a whole, and NOT to individual components such as PLATO.

C. FOR ALL STUDIES:

1. Using the course completion data, prepare another set of bar graphs like A.5 or B.3. above, but with another bar next to each of the achievement bars which shows percent in the corresponding achievement bar who

completed (did not drop out) of the course as a whole. Report the PLATO completion statistics separately only if they are of interest.

2. Prepare another set of bar graphs showing time on task for each PLATO course, ranging the data so that an average of at least 5 learners are in each bar. Note that in programs of this type, there should not be much variation in time on task. Major variations may indicate an implementation problem and should be reported. Otherwise there is no need to report this.

3. If you used a competency-based test, separate learners who mastered from those who did not. Then repeat step 5 to compare the two groups by preparing a paired-bar graph comparing time on task for each group. Range the data so there is an average of 5 learners per bar.

If you prefer, you can draw line graphs instead which compare time on task with gain or mastery status.

4. On the graph you drew in Step C.3, mark the longest, shortest and average time to mastery.

5. Separate learners who mastered from those who did not, and prepare a bar graph comparing course completion for each group, ranging the data so there is an average of 5 learners per bar.

Affective Outcomes

In any PLATO evaluation, it's common to be interested in how learners and instructors felt about their use of the system, their confidence and their role in the instructional process.

Research has shown that affective outcomes are largely independent of actual learning. Thus, asking learners how they felt about the value of a course or whether they achieved its objectives *is not* an adequate substitute for measuring learning using either of the two "program effectiveness" evaluation designs above.

Questions

Use this evaluation design when you want to find out if learners using PLATO generally feel...

- good about their experience?
- unembarrassed by wrong answers?
- in control of their learning?
- rushed?
- the system is relevant to their needs?
- the system accommodates their preferred learning style?
- confident of their ability to succeed?
- more goal-oriented?
- motivated to use the system effectively?

And, if instructors using PLATO generally feel...

- good about their experience?
- able to spend more time 1-on-1 with learners?
- in control of the instruction?
- a generally good understanding of what PLATO teaches (and does not)?
- well prepared for the changes?
- adequate time to use the system effectively?
- adequately trained to use the system effectively?
- motivated to use the system effectively?

You can answer these questions at the end of the program, or you can do a "before and after" comparison in which you measure expectations and current attitudes ("before") against actual feelings at the end of the program.

Requirements

Any PLATO implementation can gather this information. This evaluation may be combined with either of the above designs.

Procedures for Data Gathering

You can choose either an interview or a questionnaire format, and you can collect data "before and after" if you want to compare changes in attitudes due to the program, or just "after" if you are only interested in outcomes. Interviews are time-consuming to schedule and complete, but they get more in-depth (richer) information. Questionnaires are more efficient to deliver, but they get only one word or check-off answers to the questions asked (experience has shown that only a handful of highly motivated people will write even a sentence on a questionnaire).

1. Decide whether to use interview or questionnaire format.

2. If you are using questionnaire format, design at least two questionnaires by using or adapting the prototypes available from TRO, such as the ones in the Appendix. One questionnaire form should be for instructors and another for learners. You may need additional variations for training managers, lab administrators, or unique groups of learners.

3. If you are using an interview format, script the interview to include appropriate introductions, context-setting statements, and questions (contact TRO for sample scripts). Again, you may need two or more separate scripts.

4. Validate each questionnaire or script by trying it out on 3-5 respondents. Correct ambiguities and omissions so that each question means what you intended.

5. If you intend to look for changes in attitudes, administer the questionnaires or interviews before training on PLATO, or use of the system, begins. For interviews, be especially careful to stick to the exact wording in your script and record answers verbatim by hand or with a tape recorder.

6. To measure outcomes, administer the questionnaires or interviews within three days of completion of the program.

Procedures for Data Analysis

IF YOU ARE USING QUESTIONNAIRES:

Produce one bar graph for each question showing the number of people who "strongly agree," "agree," and so on. Be sure to show how many people did not answer the question. If you are studying change in attitudinal components, use a paired-bar plot with the "before" and "after" results side by side.

IF YOU ARE USING INTERVIEW PROTOCOLS:

1) copy each answer to each question verbatim.

2) group answers which are similar and note how many respondents were in each group.

3) for each question, write a summary of phrases taken from the interviews.

If you are doing a "before" and "after" comparison with an interview protocol, use a two-column format to associate comparable answers in each group.

Case Studies/Qualitative Studies

Decision-makers often have difficulty drawing appropriate conclusions from purely numerical evaluation studies. Some decision-makers have an overt distrust for quantitative studies: "figures lie and liars figure." In addition, evaluations which summarize results for an entire group require considerable effort to do, and often aren't available until long after completion of the program. In non-mastery-based classrooms where PLATO is the only instructional element, it may not be possible to use a quantitative design to study the effects of PLATO. Finally, many people find numerical evaluations too abstract to say what's really important.

A case study (qualitative) evaluation is an alternative which addresses these criticisms. In general, a qualitative evaluation can:

- Serve as a useful complement to a numerical evaluation, and "put a human face on the numbers."
- Carry more "face validity" than a numerical study, and thus have more impact with some decision-makers.
- Be prepared quickly and at low cost if the number of learners is small.
- Contain "narrow but deep" descriptions of one or more experiences with the program, rather than "wide but shallow" generalizations common of numerical evaluations.

Among the limitations of qualitative studies are:

- Observations are true only of the learners interviewed and may not be generally true.
- Data collection is less objective than numerical studies.
- Including the learner in the study may be an added motivator not available to most learners. This can lead to biased conclusions.
- Data collection and interpretation is costly unless numbers of participants is very small.

A qualitative evaluation can be used instead of or in addition to a more conventional numerical evaluation.

Questions

Use this design when you want to ask:

- What are the needs of the learner?
- What is the motivation and the goals of the learner?
- What is the learner's pattern of PLATO usage?
- How has the learner integrated PLATO with other learning experiences to achieve the class' learning goals?
- In learning from PLATO, what facilitators and obstacles were encountered?
- How has PLATO usage affected the learner's confidence?
- What have been the learning outcomes for that learner?

Requirements

A qualitative methodology can be used with any program.

Procedures for Data Gathering

You can gather data using any combination of 1-on-1 interviews, focus groups, or learner journals and writing logs.

1. Within the first few weeks of the program, identify at least 3-5 learners from each group of learners who may differ in their needs, usage patterns or outcomes.

A. FOR 1-ON-1 INTERVIEWS:

1. If possible, schedule bi-weekly, 20-minute interviews separately with each learner and instructor. Phone or send each participant a letter

explaining the purpose of the interviews, emphasizing confidentiality, and asking for participation.

If this is not possible, schedule longer interviews on a less frequent basis, any time up to one week after completion of the program.

2. Plan a series of 6-8 open-ended questions which will elicit comments on any of the questions listed above which are of interest. Test the questions for clarity and lack of bias on 3-5 "trial" interviewees.

You may wish to obtain completion/mastery status information and time on task information for each learner before each interview, and use that information as a way of prompting the learner to explain his or her experience with the system since the last interview.

3. Conduct the interviews and write down responses verbatim or tape record them. Use non-directive probes to clarify and elicit explanations. Interviews may be by phone, but are better done in person.

B. FOR FOCUS GROUPS:

1. Schedule each group of 3-5 learners and/or instructors for a 45-minute "debriefing" discussion of their experience with PLATO. You can schedule the interviews immediately at the end of the course or at both the beginning and the end if you want to compare expectations and outcomes. If the course is long (more than 2-3 months), you may wish to schedule intermediate interviews every 4-6 weeks.

2. Plan a series of 6-8 open-ended questions which will elicit comments on any of the questions listed above which are of interest. Test the questions for clarity and lack of bias on 3-5 "trial" interviewees.

3. Conduct the interviews and write down responses verbatim or tape record them. Use non-directive probes to clarify and elicit explanations. The focus group should be done with the group sitting in a circle or around a table.

C. LEARNER JOURNALS AND WRITING LOGS

1. Before the beginning of the class, make arrangements with the instructor to have the learners write down their thoughts about their experiences on PLATO as they work. A *journal* typically is a daily written or tape recorded diary. A *writing log* typically is a more lengthy open-ended writing (or tape recorded) assignment done less often.

If you and the instructor agree on a journal, have the instructor schedule 10 minutes writing time daily after completion of the time on PLATO. Learners may also write in their journals while working on PLATO.

If you and the instructor agree on a writing log, have the instructor schedule 30-minute writing (taping) periods, perhaps once per week, for the purpose of recording subjective reaction to PLATO.

If class time can't be devoted to the journals and logs, you can ask for them as out-of-class assignments, and perhaps arrange a weekly reward for those who successfully complete the assignment and turn in their work.

D. FOR ALL STUDIES:

1. If possible, when each learner has completed the program, obtain the PLATO system's data for that learner on completion/mastery status and total time on task.

Procedures for Data Analysis

Begin by transcribing all comments from the learners from your notes or tape recordings. Then group all transcribed and written comments by the "starter" questions you asked in focus groups or interviews and by topic within those questions. You're likely to discover major issues which emerged in discussion but were not directly relevant to the questions you asked. If you do, group comments related to these topics separately. If possible, note how many people made similar comments.

For each group of comments, write one or more general statements which seem to capture the meaning or significance of the comments. Follow each statement with its supporting verbatim comments.

Once the statements are written, write a narrative which describes one or more typical learner's experience in the program. Use the learner's own words as much as possible, and make clear which statements are yours and which are the learner's. State your conclusions separately at the end of each section.

Program Comparison

There is over 25 years of research showing that people can learn from computer-based instruction, and that on the whole CBI implementations achieve greater gains in less time than large-group classrooms. However, it is still true that decision makers often want to know if using PLATO is an improvement over whatever is being done now in their program. Unfortunately, meaningful evaluations of this type require carefully controlled "experimental" designs and are very difficult to do well. There are a number of reasons for this:

- "Whatever is being done now" is often imprecisely defined and varies widely from instructor to instructor, class to class and semester to semester.
- It's common to change more than one thing in a program when PLATO is introduced, so it's impossible to attribute specific changes to PLATO.
- An interpretable "experiment" has requirements which are incompatible with many program structures.
- Everyone's performance improves for a while when they know they're "part of an experiment."

Questions

Use this design when you want to ask:

- What learning gains occur under PLATO, when compared to our conventional practice?
- What percentage of learners achieve a predetermined mastery level under PLATO, when compared to our conventional practice?
- What is the longest, shortest and average time to reach mastery level under PLATO, when compared to our conventional practice?
- What are the effects of using PLATO on completion/dropout rate, when compared to our conventional practice?
- Are any of the above comparisons large enough that we can say with confidence that they are not due to chance alone?

Requirements

The comparison must use an "experimental" design in which the "treatment" group uses PLATO and the "control" group does not, but *nothing else differs between the two groups*. Note that this means that if the classroom is fixed-paced but the PLATO group is mastery model, then the comparison is between fixed-paced and mastery-model (self-paced) instruction--not PLATO vs. classroom.

General requirements are:

- Learners match the PLATO target population profile:
  - Young adult or adult
  - 3d grade minimum reading level
  - No major learning disabilities
  - General intelligence within normal limits
- Conventional and PLATO groups must run side-by-side at the same time. Running them sequentially (e.g., in consecutive semesters) is acceptable as long as all the other requirements are met (but this is often very difficult).
- All learners must come from the same general group and be randomly assigned to either treatment (e.g., by using a table of random numbers).
- It is *not* acceptable to attempt to match groups by variables such as age, sex, ethnicity, achievement level, prior schooling, schedule preference, site, work experience or personal goals. There are too many such variables to make a matching system work effectively.
- Learners may *not* select one group or the other.
- Content in both treatments must be identical. This may require constructing custom PLATO routing activities which match the classroom objectives.
- The same pre- and post-tests must be used in both treatments.
- Parallel forms of the same test must be used for pre- and post-tests.
- Tests must be administered at the beginning of the program and within three days of completion.
- The same instructor should teach both treatments. If more than one instructor must be used, each instructor should teach both

treatments and all instructors should attempt to standardize their classroom presentations and techniques. This should be confirmed by periodic classroom observations.

- All classroom treatment segments should work at the same pace, using the same syllabus.
- All classroom treatment segments should use the same assignments, texts, handouts, and other instructional materials.
- Instructor(s) using PLATO must be fully trained. If possible, evaluation should start *after* the instructor has used PLATO for one full iteration of the program (e.g., one semester). This will minimize initial enthusiasm effects and "learning curve" effects.
- There should be no contact between learners in the two treatment groups, and no opportunity for the classroom group to use PLATO.
- The PLATO implementation may use either the mastery (self-paced) or fixed-paced models. However, if the PLATO implementation uses mastery and the classroom does not, then the evaluation will be primarily a comparison of mastery and conventional practice, and results will not be attributable directly to PLATO.

Procedures for Data Gathering

1. When all learners have been selected for the program, assign each a number. Then use a table of random numbers to assign each learner to either the PLATO or conventional treatment groups. Record these assignments. Then give every learner a unique identifier to be used on all on-line and off-line activities.

If you are comparing a PLATO program with a previous iteration of the program using conventional methods, then all learners who enroll for each program iteration will receive one treatment or the other. In this case, be particularly careful that the learners do not differ over time. For example, it's likely that summer semester learners differ from fall, and that day learners differ from evening. These differences will invalidate the evaluation.

2. Administer the pre-test to all learners (if it is different from the screening test used to select learners for the program).

3. Conduct the two treatments according to their designs. Monitor each treatment continuously to verify that all the requirements are being met. Note that for the PLATO group, the requirements listed under "Mastery-Based Program Effectiveness" or "Non-Mastery-Based Program Effectiveness" also apply. If deviations from the requirements are observed, note them.

4. If you want to compare time on task between the two treatments, develop a study time log form for learners in the classroom treatment to maintain. Collect the forms regularly. Monitor the accuracy and completeness of the data entry and provide feedback to learners.

5. At the end of the program, collect time on task data for each course for each learner from the PLATO CMI system. You may also wish to collect completion/mastery data.

6. Administer the post-test to all learners within three days of when they complete the program. Since learners may be taking the test at different times, be sure to take adequate precautions to preserve test security.

Procedures for Data Analysis

A partial sample data analysis is in the Appendix.

1. Record beginning and ending raw test scores separately for each learner in each curriculum (reading, language arts, math) in each treatment group. Use a coding scheme so it's clear which score applies to each learner, curriculum and treatment.

2. If you are using an external norm-referenced test, convert each score into percentiles, stanines or grade level equivalents.

3. Compute a gain score by subtracting each pretest from each post-test score.

4. Regroup the data into not more than 10 ranges, so that the average range has at least 5 learners in it.

5. Plot three paired-bar graphs (for reading, language arts and math) comparing number of learners who scored in each range, with bars for each treatment group side-by-side.

6. If required, compute the average gain and plot a two-bar graph comparing the average gains in the two treatment groups.

7. If required, perform a *t*-test of statistical significance to compare the average gain of each group (the null hypothesis is that the groups differed by chance alone).

8. Using the completion data, prepare separate paired-bar graphs for each treatment group which compare the number of learners in each range of post-test achievement and how many completed the course.

9. On the graphs drawn in Step 5, draw a horizontal line showing gain averages for the two treatments.

10. If you gathered the data on time on task, prepare bar graphs showing time on task for each course, ranging the data so that an average of at least 5 learners are in each bar. Do this for each treatment group.

11. Separate learners who mastered from those who did not, and prepare a bar graph comparing time on task for each group, ranging the data so there is an average of 5 learners per bar. Do this separately for the two treatments, then overlay the graphs for the two treatment groups to create a paired-bar graph.

12. If the data are available, with the above bar graph, mark the longest, shortest and average time to mastery for each treatment group.

13. Prepare line graphs comparing time on task with gain or mastery status in each treatment group.

Cost Analyses

Program administrators sometimes want to cost-justify use of PLATO on the basis of reduced (or avoided) costs, increased enrollment income, or increased benefit to the host organization. This kind of cost-effectiveness justification is based on a comparison of the results of conventional and PLATO programs arrayed against the cost of each.

Do not confuse a cost-effectiveness study with a simple cost-cost comparison, in which the costs of the conventional and PLATO programs are compared without reference to the effectiveness of the programs.

Also, do not confuse cost-effectiveness with a cost-benefit study. A cost-benefit evaluation might be done for a new program to see if it is meeting its projected costs and benefits. To do a cost-benefit study, you would use the "Mastery-Based Program Effectiveness" or "Non-Mastery-Based Program Effectiveness" evaluation models, and collect appropriate cost and enrollment income data as well. If your program is justified on the basis of benefits to a host organization or the general community, then you may also be able to gather cost-avoidance data such as reduced follow-on training time, reduced equipment downtime, reduced plant or process startup, and so on. In either case, there is no comparison in a cost-benefit study with an existing program.

To do a cost-effectiveness study, you must use the "Program Comparison" design. Then you must decide if you want to do the cost analysis on the basis of reduced costs or increased enrollments. If you choose to use reduced costs, then add to the evaluation a costing model for each treatment. If you choose to use increased enrollments, then add to the evaluation design a comparison of enrollments (or retention) in each program.

Requirements

IF YOU ARE DOING A COST-EFFECTIVENESS COMPARISON EVALUATION:

- You must use the "Program Comparison" evaluation design, above.
- Number of learners in each group must be identical, or costs must be expressed as a "cost per learner."
- Cost models for each treatment group must be comparable in the way they handle fixed costs (overhead), instructor costs, learner costs (such as unemployment benefits, time off work and cost of replacement labor), and other variable costs.
- Some important benefits may not be directly convertible to dollars, but should be included in the model and reported separately from dollars. Examples include attendance, incidence of supervisory problems, incidence of safety violations, and number of learners who go on to other training or obtain employment.
- Be sure to include in your model both direct and indirect costs, and to do so in ways which are comparable for both treatment groups. For example, it would be an error to include the cost of remodeling to set up the PLATO lab (a direct cost) without charging classroom space against the conventional treatment.
- Amortization schedules applied to costs should correspond to the total period of benefit calculation.

It's often difficult to decide the "scale" of the cost model. At one extreme, the model might be limited solely to the program administrator's direct cost budget for the current program iteration. At the other extreme, the model might track costs and benefits to the employer of improved labor force capabilities over a period of five years or more. In general, the larger the "scale" of the model, the more meaningful will be the comparisons--but the harder the data will be to collect. Pick a "scale" for the model which will be of greatest use to your decision-makers, and which corresponds to your organization's mission.

IF YOU ARE DOING A COST-BENEFIT EVALUATION:

The same requirements apply but without the need for comparison between the PLATO and non-PLATO implementations. Thus, you can use either of the two "program effectiveness" evaluation designs and build a cost model which applies only to the current (PLATO) program.

Procedures for Data Gathering

1. Gather effectiveness data according to the guidelines for the evaluation model you are using.

2. If your design includes benefits to the host organization or community, contact managers, administrators or supervisors in advance and make arrangements with them on how they will collect the data, the format in which it will be recorded, and when you will contact them to retrieve it. Remember that many of these data will not be in the form of dollars.

3. Direct costs can probably be obtained from your organization's administrative accounting system. Contact your business manager or controller for data on indirect costs required by your model, amortization schedules, and for assistance in converting non-dollar data into dollars (if your evaluation calls for it).

4. If your study is a cost-effectiveness design, you will need to collect cost data "before" and "after," or separately for the two parallel treatment groups at the end of the study period.

Procedures for Data Analysis

1. Analyze program effectiveness data using the procedures included in the appropriate evaluation design.

A. FOR COST-EFFECTIVENESS EVALUATIONS:

1. Display relative cost comparisons for each treatment group by using pairs of bars, one pair for each component of your cost model.

2. Compare relative total program costs of the two treatment groups by using a stacked bar chart with two bars.

3. If the number of learners differs in the two treatment groups, recompute costs on a "per learner" basis and repeat steps (1) and (2), above.

4. Prepare a summary cost-effectiveness display which shows for each treatment group:

- Effectiveness in learning gain or mastery
- Average time on task
- Affective benefits (if included)
- Total indirect benefits to the host organization or community (if included)
- Total dollar cost per learner (using the cost model with amortization projected out over the same period as the benefits)

B. FOR COST-BENEFIT EVALUATIONS:

1. Prepare cost summaries for the PLATO implementation, using the model you designed.

2. Prepare benefit summaries for all benefits which can be converted into dollars. Use a stacked bar chart to compare these to the costs, on a total program basis, and again on a per-learner basis.

3. Prepare a table listing all non-dollar benefits.

C. FOR ALL STUDIES:

Recheck all cost models to verify that data are complete, that comparability has been preserved in data collection, that correct amortization schedules have been applied, and that all planned for indirect costs have been included.

Evaluation Procedures Summary

Often, you'll want to combine the evaluation designs. To help, here is a summary of evaluation requirements, data gathering and data analysis procedures for the evaluation designs discussed in this paper.

Requirements

Require-ments	Mastery Prog. Eff.	Non-Mastery Prog. Eff.	Affective Outcomes	Case/ Qualitative	Program Compari-son	Cost Compari-son
Learners match PLATO profile	x	x	x	x	x	x
Unlimited access	x				x	
Individual placement	x				x	
Individual routing	x				x	
Work @ own pace	x				x	
"Place out" with module tests	x	x	x	x	x	
Pass module test to go on	x	x	x	x	x	
Non-PLATO tests must be aligned	x	x			x	

Require-ments	**Mastery Prog. Eff.**	**Non-Mastery Prog. Eff.**	**Affective Outcomes**	**Case/ Qualitative**	**Program Compari-son**	**Cost Compari-son**
Non-PLATO Learning activities aligned & synchron-ized	x	x			x	
Random assignment					x	
Parallel content					x	
Post-test in 3 days of completion	x	x			x	
Implement-ation training, support & monitoring	x	x	x	x	x	x
No contact between classes					x	

Data Gathering

Require-ments	Mastery Prog. Eff.	Non-Mastery Prog. Eff.	Affective Outcomes	Case/ Qualitative	Program Compari-son	Cost Compari-son
Pre-test & prescription	x	x			x	x
Monitor implemen-tation	x	x	x	x	x	x
Mastery/ completion	x	x	x	x	x	
Time	x	x			x	x
Post-test in 3 days	x	x			x	x
Question-naires			x			
Interviews/ focus groups before, during & after			x	x		
Logs/ writing				x		
Cost data						x

Data Analysis

Require-ments	Mastery Prog. Eff.	Non-Mastery Prog. Eff.	Affective Outcomes	Case/ Qualitative	Program Compari-son	Cost Compari-son
Convert pre/post scores	x	x			x	x
Gain scores	x	x			x	x
Pre-post paired bar graph	x	x				x
Gain bar graph					x	x
t-test	x	x			x	x
Between groups comparisons					x	x
Mastery/ completion %	x	x			x	x
Time on task	x	x			x	x
Mastered vs. non-mastered time	x	x			x	x
Cost summary/ comparison						x
$ Benefit summary/ comparison						x
Non-$ benefit						x

Writing Your Evaluation Report

Here is a general outline for a typical evaluation report:

1. Program Description
    a. Location
    b. Learner characteristics and needs
    c. Program goals and structure
    d. Instructor characteristics and role in the program
2. PLATO Implementation Description
    a. Purpose of the implementation: needs and goals
    b. Installation description
    c. Implementation procedures and their success
        1) hardware/software installation & availability
        2) instructor training
        3) lab manager training
        4) ongoing support from TRO
        5) administrative support
    d. Scheduling and availability
3. Evaluation Design
    a. Description: questions and requirements
    b. Why chosen
    c. Decisions to be made based on evaluation outcomes
4. Evaluation Implementation
    a. Procedures used for data collection
    b. Observed exceptions to / problems of requirements and procedures
    c. Time of data collection

5. Evaluation data and outcomes (prepared according to the models in use)

6. Recommendations for further evaluation

7. Recommendations for decisions

    a. Program enhancements/ recommendations to instructors

    b. Program decisions/ recommendations to administrators

Appendices

Sample Learner's Questionnaire

The next page has a sample questionnaire you can adapt to evaluate your learners' assessment of their experience on PLATO.

LEARNER'S SURVEY

My PLATO LOGON ID is:	Today's Date:
My Instructor:	Location:

Part I Directions: We would like to know how you felt about your experience on the PLATO® system. For each of the statements below, please **check (✓) the box under:**

SA if you **strongly agree**
A if you **agree**
N if you **neither agree nor disagree**
D if you **disagree**
SD if you **strongly disagree**

	Question	SA 5	A 4	N 3	D 2	SD 1
1	I am able to sign on to the computer without problems.	❑	❑	❑	❑	❑
2	Getting to my lesson is easy.	❑	❑	❑	❑	❑
3	The computer is easy to use.	❑	❑	❑	❑	❑
4	I can start and stop a lesson whenever I want.	❑	❑	❑	❑	❑
5	The computer lets me answer often and not mainly just watch.	❑	❑	❑	❑	❑
6	I usually can understand what the computer teaches me, without help from my instructor.	❑	❑	❑	❑	❑
7	The computer gives me help when I need it.	❑	❑	❑	❑	❑
8	I can work at my own pace on the computer.	❑	❑	❑	❑	❑
9	I feel I'm studying what I need to on the computer	❑	❑	❑	❑	❑
10	The lessons on the computer are designed for people like me.	❑	❑	❑	❑	❑
11	When I give a wrong answer on the computer, I feel bad about myself.	❑	❑	❑	❑	❑
12	I would like more time to study on the computer.	❑	❑	❑	❑	❑
13	The computer makes me nervous.	❑	❑	❑	❑	❑
14	Working on the computer makes me feel good about myself.	❑	❑	❑	❑	❑
15	I recommend learning from the computer.	❑	❑	❑	❑	❑

Part II Directions: Please write your opinions on each the following questions:

1. What do you like **best** about learning from the computer?

2. What do you like **least** about learning from the computer?

3. How would you change the computer lessons or the way you use them?

4. What other suggestions do you have to improve any part of this course?

Sample Instructor's Questionnaire

The next pages have a sample questionnaire you could adapt for use with your instructors after they have had experience with the PLATO system.

INSTRUCTOR'S PLATO® EVALUATION

Name:	**Location:**
Class Using PLATO:	**PLATO Curricula Used:** ❑ **Reading** ❑ **Math** ❑ **Language Arts** ❑ **Other:**
Today's Date:	**How many months have you used any kind of computer for teaching?**

Part I Directions: We would like to know how you felt about your experience teaching with the PLATO® system. For each of the statements below, please **check (✓) the box under:**

SA if you **strongly agree**
A if you **agree**
N if you **neither agree nor disagree**
D if you **disagree**
SD if you **strongly disagree**

	Question	**SA** 5	**A** 4	**N** 3	**D** 2	**SD** 1
1	The PLATO course content includes what my students need to learn about the topics taught.	❑	❑	❑	❑	❑
2	The PLATO course objectives correspond to those for my course.	❑	❑	❑	❑	❑
3	The PLATO course content corresponds to the content of the standard end-of-course test we use.	❑	❑	❑	❑	❑
4	Content seemed generally free of errors and inaccuracies.	❑	❑	❑	❑	❑
5	Content was generally up-to-date.	❑	❑	❑	❑	❑
6	Quality and style of instruction was consistent throughout the curriculum.	❑	❑	❑	❑	❑
7	Students generally understood the explanations.	❑	❑	❑	❑	❑
8	There was adequate depth in exercises and tests.	❑	❑	❑	❑	❑
9	Tests, application/drill lessons, and tutorials corresponded to the objectives in the Instructor Guides.	❑	❑	❑	❑	❑

Question		SA 5	A 4	N 3	D 2	SD 1
10	Tutorials involved the students through frequent questions, answers and feedback, rather than just reading.	❑	❑	❑	❑	❑
11	Software was generally free of bugs and errors.	❑	❑	❑	❑	❑
12	All courseware used consistent keystrokes and display style.	❑	❑	❑	❑	❑
13	Color was used appropriately.	❑	❑	❑	❑	❑
14	Graphics were used appropriately.	❑	❑	❑	❑	❑
15	Screens were consistently readable.	❑	❑	❑	❑	❑
16	I was able to use student progress reports to identify students needing my attention.	❑	❑	❑	❑	❑
17	I was able to spend time in one-on-one tutoring and counseling while students used PLATO.	❑	❑	❑	❑	❑
18	I was able to make appropriate individual student assignments on the system.	❑	❑	❑	❑	❑
19	My students rarely seemed confused or "trapped" by the system.	❑	❑	❑	❑	❑
20	My students respond well to the PLATO system.	❑	❑	❑	❑	❑
21	I find working with the computer is generally a productive, rather than frustrating, experience.	❑	❑	❑	❑	❑
22	I enjoy working with the PLATO computer system.	❑	❑	❑	❑	❑
23	The PLATO system plays a useful role in my teaching.	❑	❑	❑	❑	❑
24	I was adequately trained to operate the PLATO system.	❑	❑	❑	❑	❑
25	I would like more training on how to use PLATO to best advantage in my teaching.	❑	❑	❑	❑	❑

Part II Directions: Please write your response to each question in the space provided.

1. What do you like **best** about teaching with the PLATO computer?

2. What do you like **least** about teaching with the PLATO computer?

3. How would you change the PLATO lessons?

4. What suggestions do you have to improve the way you use the PLATO system?

5. What other comments or suggestions do you have on the PLATO system or this course?

Example Data Analysis

The examples below show two kinds of data analysis used in these evaluation designs.

First, here is an example of how to tabulate the data from an external test. This is a competency-based test, so the numbers are absolute scores rather than grade level equivalents. The table shows data for a Program Comparison design comparing a classroom (ILT) to a PLATO (CAI) program. The ILT group had complete data for 45 learners, and the CAI group had complete data for 36 learners. However, only the data for 15 learners are shown here.

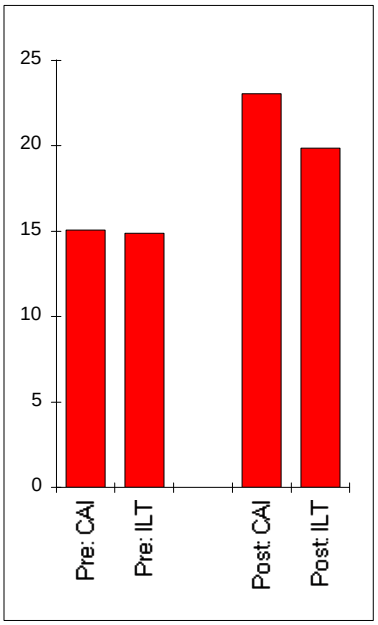
Average Math Test Gains: Instructor-Led (ILT) vs. PLATO (CAI)

Learner #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Pretest:ILT	6	8	9	9	9	10	10	10	11	11	12	12	12	13	13
Posttest:ILT	21	27	12	20	22	10	24	15	13	23	23	11	22	11	13
Gain:ILT	15	19	3	11	13	0	14	5	2	12	11	-1	10	-2	0
Pretest:PLATO	6	7	7	9	10	10	10	11	12	12	13	13	13	13	14
Posttest:PLATO	16	21	24	14	10	12	19	25	20	32	17	27	28	33	12
Gain:PLATO	10	14	17	5	0	2	9	14	8	20	4	14	15	20	-2

Here are the averages (computed for the entire data set):

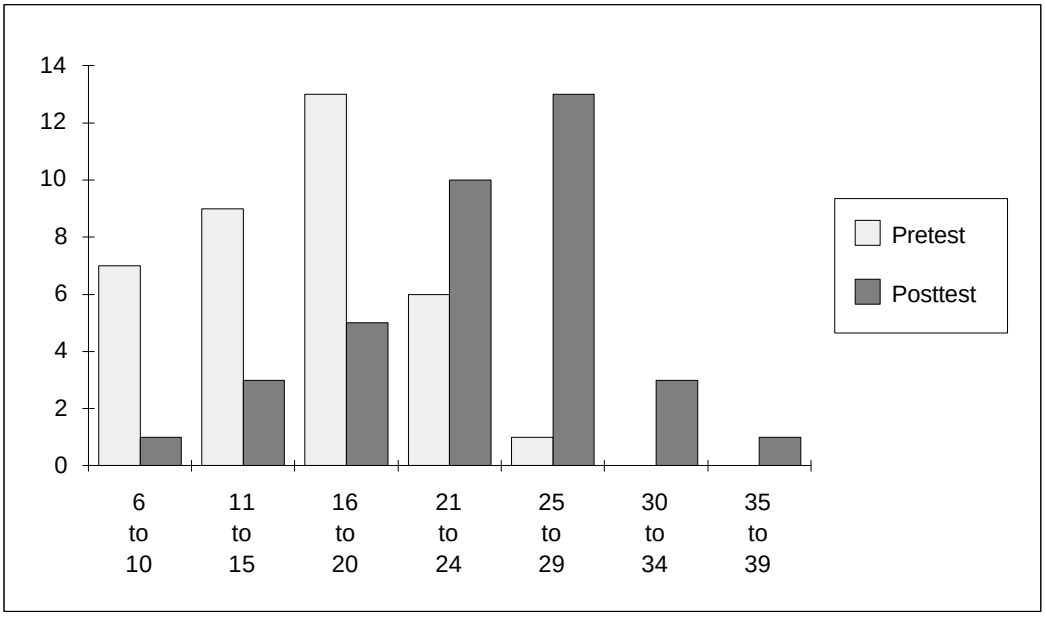
Pre: CAI	Pre: ILT		Post: CAI	Post: ILT		Avg Gain:CAI	7.9722222
15.055556	14.888889		23.027778	19.866667		Avg.Gain:ILT	4.9777778

The averages may be charted like this:



The kind of analysis most commonly recommended in the evaluation designs calls for comparisons of paired series of bars. This is an example, based on the data above. The pretest and post-test scores (not gains or averages) have been grouped by 5's and then are plotted together. This is more informative than simply showing the averages as above.

	6 to 10	11 to 15	16 to 20	21 to 24	25 to 29	30 to 34	35 to 39
Pretest	7	9	13	6	1	0	0
Posttest	1	3	5	10	13	3	1



Further Reading

Herman, Joan (ed.), *Program Evaluation Kit*, 2nd Ed. Newbury Park, CA: Sage Publications, Inc., 1991.

This is a very practical set of booklets on evaluation. Order singly or the entire set by mail from:

Sage Publications, Inc.

P.O. Box 5084

Newbury Park, CA 91359-9924.

Phone: (805) 499-0721 Fax: (805) 499-0871.

1991 Series price was under $80. Individual titles and 1991 prices are:

Vol. 1: Evaluator's Handbook ($12.95)

Vol. 2: How to Focus an Evaluation ($8.95)

Vol. 3: How to Design a Program Evaluation ($9.95)

Vol. 4: How to use Qualitative Methods in Evaluation ($9.95)

Vol. 5: How to Assess Program Implementation ($9.95)

Vol. 6: How to Measure Attitudes (($9.95)

Vol. 7: How to Measure Performance and Use Tests ($9.95)

Vol. 8: How to Analyze Data ($12.95)

Vol. 9: How to Communicate Evaluation Findings ($8.95)

