DOCUMENT RESUME

ED 303 156 IR 013 626

AUTHOR	Baker, Eva L.
TITLE	Sensitive Technology Assessment of ACOT.
PUB DATE	Apr 88
NOTE	9p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 5-9, 1988).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Academic Achievement; Educational Research; *Educational Technology; Elementary Secondary Education; *Evaluation Criteria; *Instructional Innovation; Microcomputers; Models; *Program Effectiveness; Program Evaluation; *Research Methodology
IDENTIFIERS	*Apple Classroom of Tomorrow

ABSTRACT

This paper explores the ideas and the model underlying the evaluation of the Apple Classroom of Tomorrow project (ACOT), a 2-year-old research and development project incorporating at least seven different grade levels which is located in five different school sites in four states. The major features of ACOT are identified as the ideas of computer saturation and local site development. Of particular concern in this paper is the evaluation of the technology within the context of the usual goal-focused educational assessment. It is suggested that the assessment of the technology assists in the identification of appropriate goals through a process entitled technology push, which allows for outcomes to be empirically identified as desired consequences of the technology's implementation. A model for the assessment is provided that attempts to incorporate evaluation features central to the schools in which ACOT has been developed, i.e., educational goals, processes, and outcomes. Finally, specific evaluation questions are formulated for the comparison of students at different ACOT sites, and specific achievement measurements are suggested. The text is supplemented by two figures, and three references are provided. (EW)

Sensitive Technology Assessment of ACOT

Eva L. Baker

Center for Technology Assessment
UCLA Graduate School of Education

Paper presented at the Annual Meetings of the
American Educational Research Association
and the
International Association for Computing in Education

for the Symposium
"Natural Experiments in Computer Saturation:
The Apple Classroom of Tomorrow"

April 5 & 6, 1988, New Orleans.

Sensitive Technology Assessment of ACOT

Eva L. Baker
Center for the Study of Evaluation
UCLA Graduate School of Education

The purpose of this paper is to describe the ideas and the model underlying the evaluation of the Apple Classroom of Tomorrow project (ACOT). ACOT is a two-year-old R&D project located in five different school sites. The major features of ACOT are the ideas: 1) of computer saturation; and 2) of local site development. Computer saturation operationally means two computers for each participating student and teacher: one for school and one for home. Local development means that the site staff of teachers and administrators, with technical assistance, decide how best to employ such technological largesse. The ACOT sites are in four states, in 14 classrooms, and incorporate at least seven different grade levels. No single site is comparable to any other ACOT site in socioeconomic status or in grade level, and certainly not in curriculum focus.

Thus, the evaluation of ACOT provides many of the same problems confronting evaluation of distributed programs—problems of comparability, problems of knowing what the actual changes are, problems of interpreting data. Some of these problems have been described in an earlier position paper (Baker, 1987). Key issues for evaluation of innovations were identified, including audiences and their expectations; the relationship between evaluation findings and scientific findings; evidence; and the roles of those who are evaluated in shaping and interpreting the agenda. But ACOT, and, we expect, other similar technologically based innovations, present other challenges and these will be identified in the next section.

Technology Assessment in Schools

Evaluation approaches to technology must differ from typical school based evaluation because of the fundamental differences in the evolution of social and educational programs and the development and implementation of technology. Most educational programs are created to meet identified needs, and much of the experience in the evaluation field is with requirements-driven programs and products.

Examples of commonly evaluated innovations include preschool programs, specific reading texts, science curriculum, teacher training options, and procedures for involving parents in school planning. Underlying the implementation of these programs and products may be very specific goals: to help children read better or learn more science; to

---

This paper was partially supported by a grant from Apple Computer; however, the opinions expressed herein do not necessarily reflect the position or policy of Apple Computer.

have teachers teach more effectively; and to develop better school plans. When goals are stated (such as in reading materials) or strongly implied in an innovation, requirements for evaluation can be formulated rather directly. The critical issue of comparison is easy, for we can compare the innovation to present practice or to another alternative against the standard set by the goal. The innovation gets assessed in terms of its relative or absolute contribution to the attainment of the goals.

Goal-Focused vs. Technology-Focused Assessment

How does assessing a goal-focused evaluation map onto evaluating a new technology such as computers? It depends. Sometimes the technology is linked directly to a particular set of explicit goals. For instance, when the technology functions as a delivery system for a set of courseware that teaches algebra, our evaluation can proceed much as the general precepts suggest. We assess the technology's contribution to meeting the stated goals of algebra proficiency, and separate, as appropriate, the courseware and hardware contribution.

But all technological innovations are not focused on clear goals. The reason for providing computer support for students is often unconnected to a straightforward goal related to student learning. Rather, computer acquisition in schools is sometimes motivated by a more general desire to improve educational quality, impelled by a set of beliefs that the technology will somehow affect the quality of students' learning. Such belief-driven innovations may specify no goals, but may prescribe instead some minimum set of processes that constrain the innovation's use. Constraints might require that the computer is to be used by the teachers only for instructional management purposes, or by the students in a learning center or laboratory setup, or they might require that each student will share a computer with another classmate for word processing.

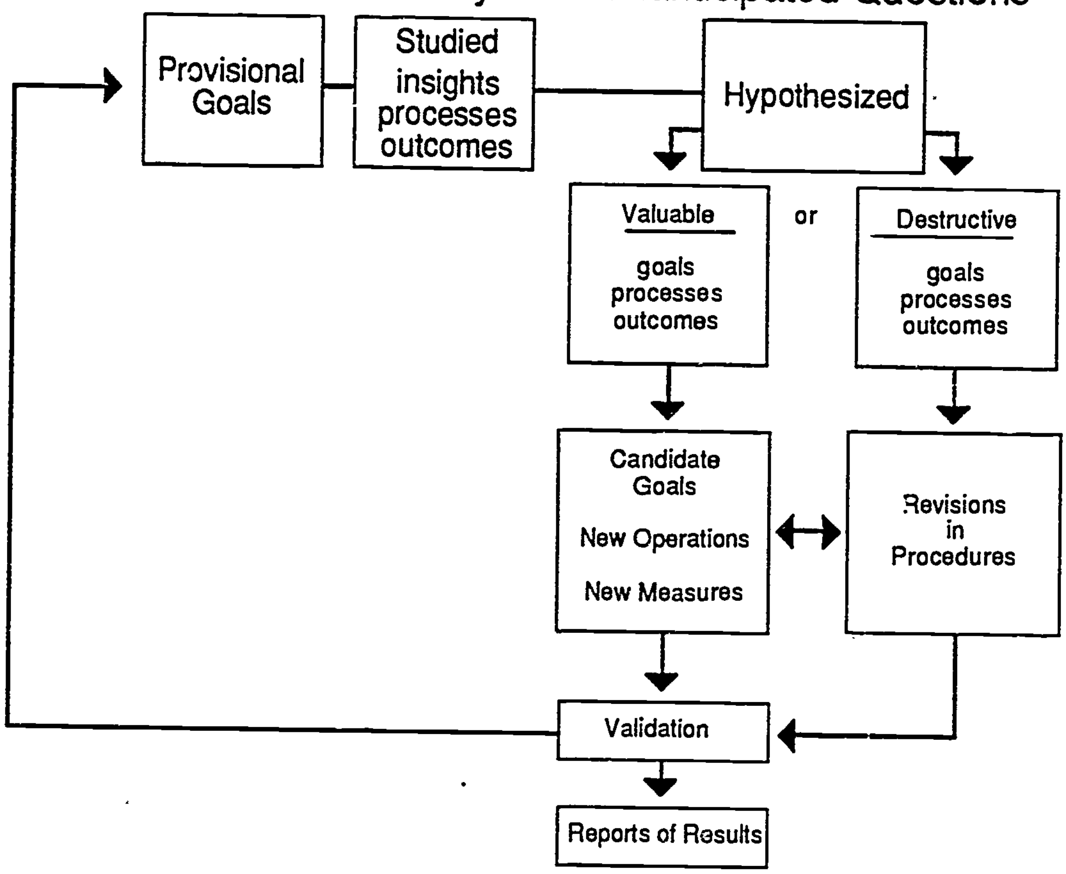
Finding Good Ripples

In other cases, even the processes are left unspecified, and we frequently drop the technology into an environment and watch what happens. Here the technology functions much less like an educational program, and much more like a utility, such as the telephone. Forecasting its potential uses is difficult. But we can watch the ripples. The assessment focus shifts from the comparative attainment of specific goals. Instead, the assessment focuses on the exploration of what processes get tried and what outcomes are possible. Converting the possible, desirable outcomes into tentative, new goals for the technology may be the most important function of the technology assessment process. Thus, the assessment of technology explicitly helps to identify appropriate goals, goals that can in the future be used as markers for all the general evaluation precepts described in the first section of this paper. This process, called technology push (Glennan, 1968), allows for outcomes to be empirically identified as desired consequences of the technology's implementation—good ripples. This feature, unique to technology evaluation, requires an extra conceptual step in our evaluation model—finding goals. It further requires that a period of protected exploration precede the implementation of any formal evaluation. This period is needed not only because goal conversion is a necessary step for any formal evaluation, but also because new technology takes some getting used to (mostly for adults, however).

Technology assessment has more than a single extra step. In point of fact, there is a fuzzy interplay between goal identification, description of processes, and assessment of outcomes (Figure 1). When the exploratory assessment period draws to an end, someone, probably a set of educational policymakers, needs to select among empirically developed outcomes and validate them as legitimate goals—goals that should be the focus of classroom activity and should subsequently be assessed according to more traditional evaluation precepts.

Figure 1

STAR Model for Analysis of Unanticipated Questions



Embracing the Heisenberg Principle

The dynamics of the interplay among observation, empirical outcome identification, and the conversion and validation of goals leads us to think much less like experimental psychologists and more like natural scientists who are trying to describe or map phenomena. One of the great insights in this process was derived from Heisenberg, who posited that measurement (or scientific observation) almost always had a detectable effect on the phenomenon being studied. Demonstration of this principle has not had much impact on behavioral science at all except to encourage people to redouble their efforts to control extraneous variables.

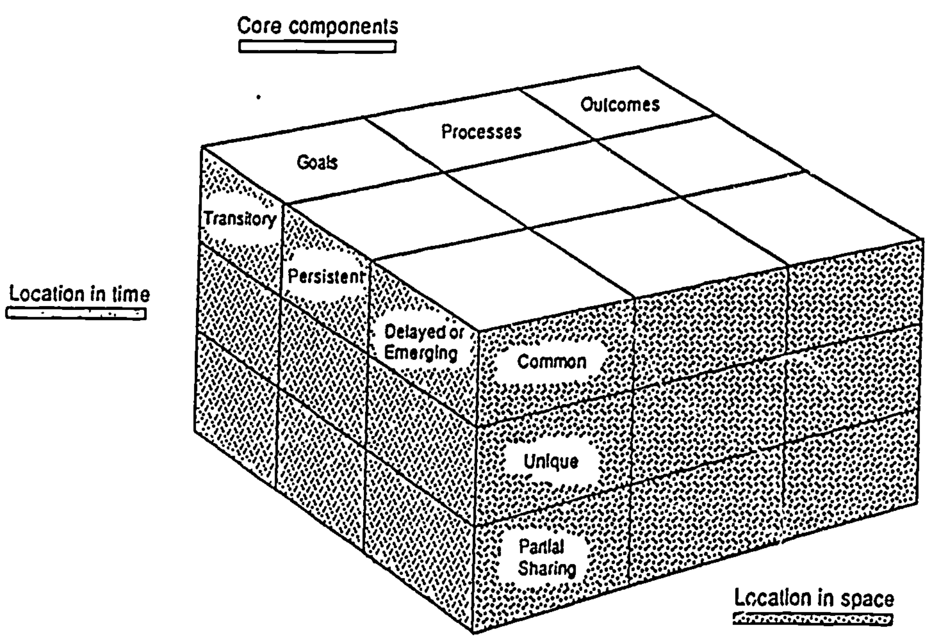
But let's argue the alternative—in two parts. In the exploratory phase of technology assessment, the fact of assessment, the careful observation and helpful interaction, might very well encourage individuals who might otherwise let their resistance, anxiety, or inexperience with technology result in cobwebbed keyboards. By having the evaluation activity descriptive, non-threatening, and very visible, the evaluation itself could stimulate more energetic exploration of technology options—just the opposite of typically controlled studies. Second, in the phase of assessment following the conversion and validation of goals, the procedures used to evaluate classroom process and student progress should be embodied in a usable system, such as a database, through which the participants as well as the evaluators can cooperate. Creating a noticeable, technology-based system has a slew of benefits. First, we demonstrate one more useful function for the technology. Second, the system concretizes interest in learning outcomes, and mobilizes involvement if participant's don't like the form in which they are measured (the usual object of disdain is multiple choice tests). So making the assessment part of the solution rather than part of the problem can be uniquely accomplished by computer technology and should be as well.

The STAR Model

Evaluators are wont to create flow charts describing what features they address as they conduct their work and call them models. We have attempted to mediate our task of developing a comprehensive evaluation plan by positing a model of the critical features of the ACOT implementation, at least to this point in 1988. Because ACOT is in some ways such a noisy environment, we thought it best to abstract our view of its essence, and then design approaches to evaluation around it. The STAR model attempts to capture the central stuff of schools: educational goals, processes, and outcomes (Figure 2).

Figure 2

STARFRAME
Framework for ACOT Phenomena



6

The first major dimension for focus is based upon the issues identified as critical in technology assessment: the emergent, adaptive, and potentially diffuse development of goals, processes, and outcomes stimulated by the innovation itself. Adding the time dimension to our fundamental classification of goals, processes, and outcomes seems to do this for us. The second dimension is one that is especial relevant to ACOT's history of development and the larger goals of this experiment: the distributed, idiosyncratic implementation of technology.

Clearly, these dimensions are not sufficient to represent fully the complexity of evaluation processes, nor to provide the detail desired by potential critics or users of the STAR approach, of ACOT, or of other new technology implementations. What we need is a way to capture presently conceived, or add or delete, based on our developing understanding, additional attributes, experiences, or data. We hope to permit the user of our model to sort through additional features or dimensions relevant to any side of the cube, e.g., core components (goals, processes, outcomes), relevant to a slice, e.g., goals, or relevant to a cell, transitory, common goals.

We plan to think about what additional dimensions or features might be included. For example, when thinking about goals we could assess them in terms of some of the following dimensions:

Goal Dimensions

- Implicit to explicit
- Challenging to comfortable
- Few to many
- Targeted to diffuse
- Critical to optional
- Goal instigators or sharers: Students, teachers, ACOT districts, parents

There is no intention to fill in a factorial set of dimensions on our part. Certain areas might be rich and have layers of information, features, or experiences, while others will be relatively empty.

We would expect that individual sites would want to characterize or to plan their programs differently. Clearly, we see this model, if implemented in database form, to be a rich resource for designers and assessors of significant technological innovations. Its utility will accrue, if for no other reason than school boards and the public will want to estimate the benefits of large investments.

The Acot Evaluation Itself

Characteristics of the ACOT evaluation:

- Interactive participation in the evaluation study by ACOT participants, collateral university researchers, and UCLA staff to develop a credible, adaptive set of assessment plans, procedures, and reports assessing the ACOT experiments;
- A phased implementation designed to conform to the rhythms of site by site implementation;

- A focus on exploring the utility of developing into a multiuser evaluation system, for future coordinate implementation with new technology adoption in local districts.

Relating STAR to our ACOT Evaluation

Our overall evaluation questions include the following:

- What are the effects of ACOT on students?
- How does ACOT influence the organization and delivery of instruction?
- How does ACOT effect teachers?
- What are ACOT effects on the family of ACOT students?
- What other unintended effects, either positive or negative, may be attributed to ACOT?

Our 1987-88 work is focused on the effects of ACOT on student outcomes, because they are necessary to satisfy the expectations of our particular sets of clients, and the remainder of the paper will deal with this topic. A major methodological problem relates to the site distribution of ACOT. Because ACOT's early development principles emphasized local site control, and because ACOT's distribution across grade level is confounded with site, we will be hard pressed to make general statements about ACOT effects. Nonetheless, we adopted certain approaches designed to address the idiosyncrasies of sites directly.

For each site and grade level within sites, we are interested in knowing the comparative impact of ACOT on four sets of students:

1. ACOT students over time, as they go up the grade structure of schools;

2. Entering ACOT students taught subsequently by an increasingly experienced ACOT teacher team, for instance 5th grade students in 1988 compared to 5th grade students in 1989;

3. Comparable students not participating in ACOT in the same school;

4. Comparable students not participating in ACOT in another school in the district.

In order to attempt to deal with longitudinal comparisons and to have a framework for making cross-site assessments, our evaluation plan alls for a variety of outcome measures. This discussion will focus on only the achievement measurement problem. Our first data points will come from district- and site-provided evidence about student performance. This evidence may take the form of existing district test results or other measures that are site or grade level specific. Second, we plan to administer to all sampled ACOT and non-ACOT students a standardized achievement test battery, either the Iowa Test of Basic Skills or the Iowa Test of Educational Development, as appropriate to grade level. We do this even though as evaluators we are skeptical that the ACOT intervention will have demonstrable effects on such measures. However, we have good reasons. First, we will ask teachers of all sampled classrooms to respond to the test scales to give us a sense of the levels at which components of the tests have received instructional attention in their classroom. These data will give us the ability to chart differential goal emphasis on such measures. Secondly, the teachers' reported emphases will permit us to control instructional exposure in our analyses of student performance and provide a somewhat more valid assessment of classroom performance. Third, our use of Iowa Tests will help us understand better the longitudinal data because of the relative strength of vertical equating of the tests from grade to grade. Fourth, we will attempt to use at least some subscales of the Iowa Tests as anchor tests to permit equating across sites to other disparate measures.

Written Composition

The second major achievement measure we are using derives from our International Education Association (IEA) study of student written composition (Baker, 1987). With selection and adaptation by ACOT site personnel in process, we plan to ask students to write narrative, descriptive, and expository prose at all participating grade levels. Our prior IEA experience has provided us with validated scoring schemes, efficient training procedures, and national and (soon) international data for comparisons. We also will ask at least some students to write about the ACOT experience itself, and based on content analyses, have a validation base for other affective data we are collecting. UCLA, with federal support, is presently conducting research and adapting the scoring approaches to measure deep content knowledge, particularly in history. If these efforts continue to be successful, subject matter content measures could also be employed at appropriate ACOT sites.

Conclusion

Our goal is to assess ACOT credibly and fairly using as strong a measurement base as possible. We will be conducting a series of validity studies to permit us to improve our work. We hope that one outcome will be the development of a set of feasible measures and strategies that can be used to assess other technology-based interventions in the future.

References

Baker, E.L. (1988). *Sensitive, sensible technology assessment.* Los Angeles: UCLA Center for the Study of Evaluation.

Baker, E.L. (1988). *U.S. Study of Written Composition*: 1988. Los Angeles: UCLA Center for the Study of Evaluation.

Glennan, T.K. (1967). Issues in the choice of development policies. In T. Manschak, T..K. Glennan, & R. Summers (Eds.), *Strategies for research and development.* New York: Springer-Valdez.