COMPREHENSIVE EVALUATION OF COMPUTER COURSEWARE: GETTING BACK TO THE BASICS

Kenneth A. Sirotnik

CSE Report No. 237

1984

CENTER FOR THE STUDY OF EVALUATION

Graduate School of Education

University of California, Los Angeles

The project presented or reported herein was supported pursuant to a grant from the National Institute of Education, Department of Education. However, the opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by the National Institute of Education should be inferred.

It should not be necessary to write this paper. Computer courseware is simply part of a schooling curriculum. Evaluating courseware should be a natural part of curriculum evaluation. The principles and practices of curriculum evaluation have been developed and refined for nearly four decades ever since the seminal work of Ralph Tyler in the late forties. Nothing terribly new will be found in this report that was not either explicit in or implied by Tyler (1947) or that has not appeared in modernized and expanded work on curriculum theory (e.g., Goodlad, et al., 1979).

Yet it seems that we are rather easily seduced by the promise of educational innovations--microcomputers being the most recent example. And, in our attempt to evaluate the promise, our thoughts turn more to surface-level, technological issues rather than to deeper meanings that we have known for some time to constitute the basic questions of curriculum and instruction.

Consider the advent of educational television; we became preoccupied with questions like: How many? Where should they be located? What are the trade-offs between black-and-white versus color? What is(are) the optimal: Number of minutes per program; ratio of talk to action; ratio of cartoon to real-life content; tonal qualities for attention-getting and maintenance? When learning machines came along, we asked: How many? Where should they be located? How small can they be? How much text should there be per frame? What are the trade-offs between linear versus branching programs? What are appropriate forms of reinforcement for correct answers and feedback for errors? What is the optimal ratio of active and passive student interface with the machine? Is it best suited for drill and practice versus higher level thinking and learning?

Clearly there is a generic set of questions here as can be inferred from the ones now being asked about computers and courseware: How many? Where should they be located? Relative advantages of color and graphics? Amount of text per screen? Appropriate use of sound and voice simulation? Appropriate forms of corrective feedback and reinforcement? Optimal ratios of student versus program control? Teaching facts and comprehension versus higher-order, problem-solving modes? These questions, and others like them, raise important issues indeed. For ease of discussion here, I will include them all in a general category labelled technology. Perhaps the single, most important representative of this category is the following question: What does the technology do that could not have been done at least as efficiently and effectively with ordinarily available learning resources (teachers, peers, paper and pencil, books, manipulatives, etc)? Many documents are now available that suggest evaluative criteria for courseware technology and I will not belabor the issue any further here.[1]

But, as hinted at above, there are two other categories of questions that in many ways transcent the particulars of any given technology. Again for purposes of discussion, I will label these the categories of curriculum and assimilation. By assimilation, I mean the process by which schools innovate; that is, in this case, the ways in which a new technology becomes part of the everyday work life of teachers and students. Unfortunately, the track record of districts and schools in innovation and change generally, and the assimilation of technology specifically, is not good. With regard to microcomputers and accompanying software, the cynic might well ask: Will these testimonials to human genius take their place on shelves along side their counterparts of past eras (teaching machines, educational TV, and the like) only to collect the dust of innovative

non-events?

The optimistic evaluator must ask: How has the learning environment (human and material) been modified to receive and exploit (in the best sense) the full potential of computer courseware? To pursue this issue further is beyond the scope of this paper. Obviously, from an evaluative standpoint, it would be foolish to hold any particular piece of courseware accountable for the larger issue of assimilation. However, since the viability of computer courseware rests upon the adequate resolution of this issue, it deserves special mention in any evaluation framework.[2] I would now like to turn to the main evaluative thrust of the present paper--curriculum.

Hopefully in what follows, the emerging evaluative questions, should they strike you as old and obvious, will strike you as essential and preemptive of those that might be leveled at the technology (e.g., micros) and its by-products (e.g., courseware). However, given the immediate purpose of this paper, these issues will be discussed in the context of computer courseware developed for classroom use.

Let's begin with content and what it is not: CONTENT IS NOT THE CURRICULUM. A comprehensive view of curriculum and any aspect of it (e.g., computer courseware) requires explicit acknowledgement and consideration of all relevant elements of the teaching and learning experience.

The fact that content per se is part of but not the whole of curriculum was implicit in Tyler's (1947) rationale. Since then, various attempts have been made to sort out the facets of curriculum as a multidimensional construct. One that has seen a good deal of theoretical and practical development is Goodlad's (1979) notion of curricular commonplaces. This idea has enjoyed considerable success both in framing

curriculum inquiry (Goodlad, Klein, & Tye, 1979) and generating relevant data for the study and assessment of schooling (Klein, Tye, & Wright, 1979; Sirotnik & Oakes, 1981).

The list of curriculum facets that can be seen as commonplace to all organizations of teaching and learning experiences includes at least the following:

1. Goals/Objectives--statements of intended teaching and learning specific enough to convey at least the relevant content and expected behaviors.

2. Content--substantive strands and topics comprising the "stuff" of teaching and learning.

3. Strategies--instructional methods or processes designed to promote teaching and learning, e.g., use of open-ended questions, group discussions, lecture, deductive/inductive approaches, etc.

4. Activities--events, tasks, etc. designed to engage teachers and learners, e.g., reading, writing, listening, practicing, role playing/simulating, etc.

5. People--human resources available to facilitate teaching and learning, e.g., teachers, aides, student peers, etc.

6. Materials--physical resources designed to facilitate teaching and learning, e.g., pencils, paper, books, learning kits, manipulatives, calculators, computers, television, etc.

7. Grouping--ways in which human resources are organized for teaching and learning, e.g., total class, individual seat work, cooperative learning groups, etc.

8. Time--allocation and use of time in teaching and learning.

9. Space--ways in which classroom areas are organized for teaching and learning.

10. Assessment--determining, collecting, and interpreting information for describing and judging the effectiveness of the teaching-learning process and for facilitating decision-making and action-taking towards the improvement of that process.[3]

But these commonplaces alone do not provide an adequate framework for making curricular judgements. They suggest appropriate types of descriptive information but lack the bases for evaluation. The commonplaces answer "What is?" questions, but do not automatically address, "Why and what ought to be?" questions. The needed evaluative screens--what I will call here educational values and beliefs--permeate all curriculum inquiry. It is simply a question of making them explicit along with, hopefully, some working consensus among those concerned about why these "oughts" are important.[4]

I have a list of educational values and beliefs that I will share with you, more for the purpose of illustration than indoctrination. Many "oughts" on this list will be familiar to you since they appear in most formal curriculum documents at state and local levels. Most are implicit in Tyler's (1947) discussion of the sources and criteria for, and the organization and evaluation of, learning experiences. Many are implicit in the Goodlad, Klein and Tye (1979) dimension of "qualitative factors" which they suggested as evaluative screens for curriculum commonplaces. All are compatible with my own orientations and experiences as a student, teacher, and parent, and as an educational researcher in collaboration with my colleagues on A Study of Schooling (see Goodlad, 1983) and, more recently, on projects of the Laboratory in School and Community Education.[5]

Consider, then, the following list of keywords, each intended to represent a constellation of educational values and beliefs:

1. Equity--equal access to the curriculum (content, teaching practices, time, etc.) regardless of race, sex, religion, etc. or any correlates thereof (e.g., social economic status).

2. Experience--building upon concrete, real-life events, feelings, and meanings in the empirical world of teachers and students.[6]

3. Critical Thinking/Problem Solving--going beyond the necessary facts and comprehension levels of cognitive processes, questioning knowledge, and using higher order processes such as analyzing, synthesizing, proving, applying, abstracting, and evaluating.

4. Discovery/Creativity--freedom to explore knowledge, think divergently, invent, imagine, and so forth.

5. Proactivity--deliberate involvement of students in their own learning such that they become active, non-passive, and non-reactive decision-makers.

6. Integration--treating knowledge "ecologically," not as discrete, unrelated bits of information, but as parts contributing to a whole.

7. Variety--deliberate use of different instructional activities, materials, grouping techniques, etc. in contrast to an over reliance on only one teaching-learning configuration (e.g., teacher lecturing to the total class).

8. Individual Variability--recognizing individual differences in ability, learning styles, attitudes, interests, etc. as assets rather than liabilities and adjusting/adapting/modifying curricular elements to accommodate these differences.

9. Socialization--humanizing knowledge through exploring why and how it is not independent of its sociocultural and political context.

10. Personalization--humanizing knowledge through exploring personal meanings, sentiments, interests, and future aspirations.

Clearly, these categories are not mutually exclusive. Some might even be combined with no loss to the usefulness of the framework; for example, critical thinking, problem-solving, discovery, and creativity might all be combined into a general category labeled inquiry. Others need to remain separate yet clearly interact in fundamental ways (e.g., considerations of equity render ability tracking indefensible as a school policy for dealing with individual variability).

Nevertheless, whether these or some others, some list of values/beliefs must be made explicit as a frame of reference for evaluating curriculum. A convenient way to map this evaluative task is to form the matrix of questions that naturally emerges by crossing an educational values and

beliefs list with the list of curriculum commonplaces. (See Figure 1.) Each cell represents the obvious set of questions that are generated by the interaction of the value/belief represented in a given row and the curriculum commonplace of a given column. For example, following through with a commitment to dealing effectively with row eight of the matrx requires asking questions like: Do the topics, activities and teaching strategies accommodate the different learning styles, abilities, etc. of the students? Are appropriate material and human resources available and used for accommodating these differences? Are students taught en masse, at the same time, and in the same place or are allowances made depending upon individual differences (using small groups, variable pacing, learning centers, etc.)? Is testing primarily summative for the purposes of uniform grading or formative for the purposes of diagnosis and facilitating individual learning progress?

The point of imposing this kind of comprehensive, evaluative screen on some computer courseware program should be obvious. Courseware does not exist independently of curriculum. It contains and/or addresses-or should address--all curriculum commonplaces. As consumers, we must evaluate the courseware accordingly. We must resist the temptation to be sold only by the flashy novelty of a technological invention and demand some understanding and judgement regarding how the invention fits into the desired curricular scheme of things.

The values and beliefs in Figure 1, for example, lead one to ask questions like: Is the courseware biased with respect to one or more demographically defined student subgroups? Can the courseware tap into real-life student experiences? Does the courseware address higher-level

CURRICUM COMMONPLACES

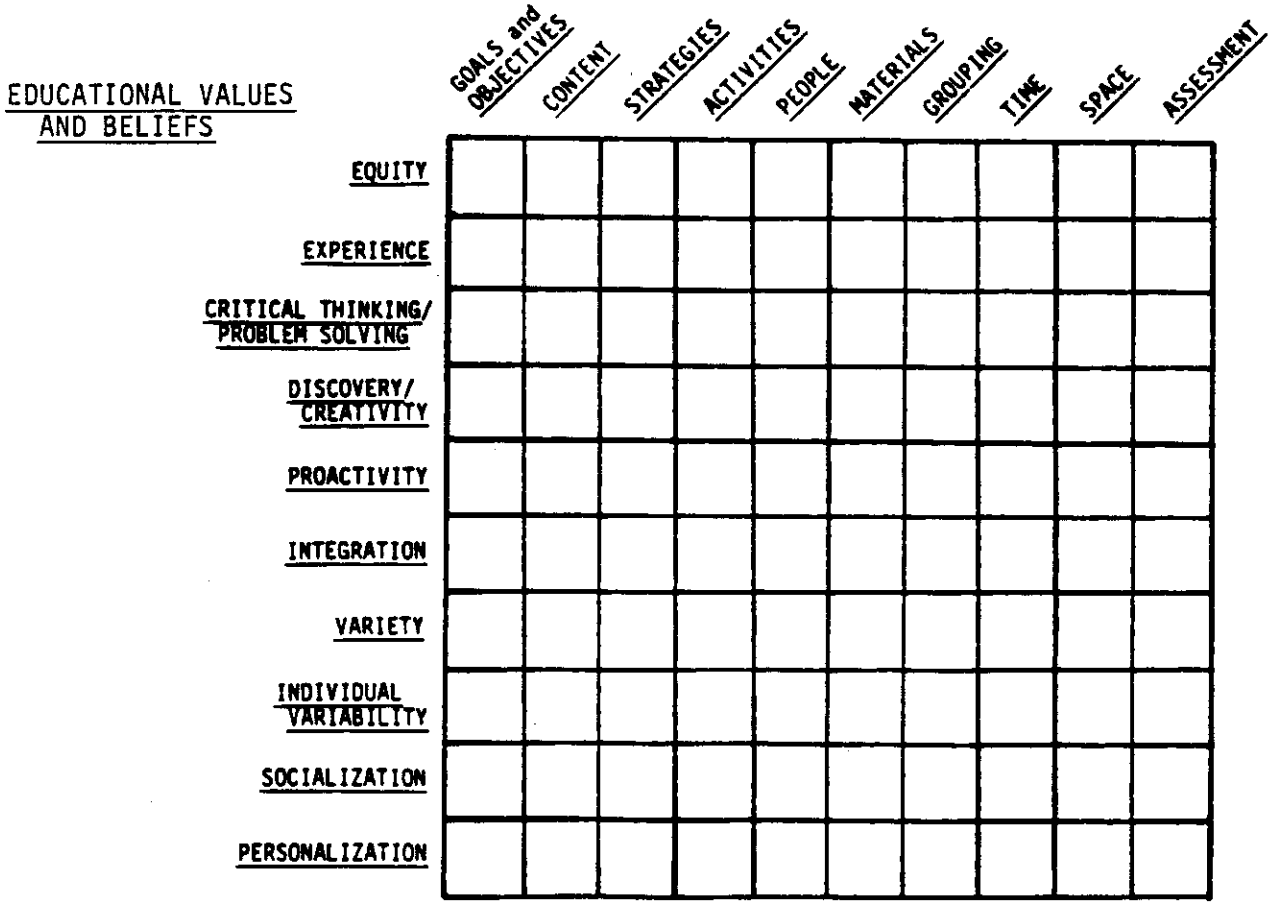

Figure 1

An Evaluative Matrix
for Describing and Judging
Curriculum

cognitive skills and processes? Is discovery learning, the exploration of concepts, and inventing new concepts encouraged by the courseware? Is it clear how the courseware is an integral part of the larger curriculum? Are individual differences in ability, learnings styles, and so forth accommodated by the courseware? Is it clear how the courseware is but one instructional vehicle among many, i.e., how it can be inferfaced with teachers, peers, ordinary materials (pencil, paper, manipulatives, etc.), learning center activities, etc.? Is the learner treated by the courseware as a passive recipient of knowledge or as an actively engaged learner and decision-maker?

These kinds of questions and/or others like them must be addressed when evaluating courseware or any significant aspect of curriculum. In so doing, it mitigates against simplistic evaluations of courseware like: "Oh, that's just a drill and practice worksheet or textbook simulator" or "Look how wonderfully this program simulates human intelligence." It is almost as though the labels "drill and practice" and "artificial intelligence" carry with them self-evident properties of "bad" and "good" respectively. Clearly, however, drill and practice courseware and programs such as LOGO can be either useful or useless depending upon how their use addresses the issues and questions suggested by a matrix such as the one in Figure 1.

In summary, it should be understood that I am <u>not</u> suggesting that any particular piece of courseware be held accountable <u>in and of itself</u> for each and every cell of a values/beliefs-by-commonplace matrix. Rather, the suggestion is that curriculum must be held accountable in this way and, therefore, so must educational software. To put it another way,

checklist-type evaluations of courseware stripped of their instructional context will be insufficient to guide selection. Certainly the information collected by these checklist evaluation techniques is useful, but particularly as it is brought to bear upon, and revised in accordance with, the intended classroom curriculum. Ultimately, thoughtful consideration of the tough questions of curriculum inquiry must be imposed upon courseware as it is used in the specific educational setting.

Footnotes

1. See, for example, the checklists in reports by Edwards (1984), Marshall (1984), Merrill (1984), Hively (1983), and Van Buskirk (1983); see also the guides published by (1) the California Library Media Consortium for Classroom Evaluation of Microcomputer Courseware, (2) The Computing Teacher, (3) the Educational Products and Information Exchange, and (4) MicroSIFT (The Computer Technology Program, Northwest Regional Educational Laboratory).

2. Many readings currently exist in the area of school innovation and change, generally, and the assimilation of technology, specifically. Examples are: Oettinger (1969); Sarason (1971); Goodlad (1975); Heckman, Oakes and Sirotnik (1983); Mayer (1984); Oakes and Schneider (1984).

3. All these curriculum commonplaces deserve considerable elaboration but that is beyond the scope of this paper. Nevertheless, assessment deserves special mention because of the popular tendency to equate it only with student achievement testing. Certainly some information is conveyed through an accounting of items answered correctly, especially on a test designed to measure course objectives (i.e., criterion-referenced testing). But much more is possible and desirable, for example:

   - exploring error patterns as in answer-until-correct formats for multiple-choice items (Wilcox, 1984).
   - developing testing strategies commensurate with various learning styles suggested by recent work in cognitive psychology (e.g., Glaser, 1981; Mayer, 1984).
   - thinking of assessment as formative (vs. summative) and incorporating routines being recently suggested in the areas of diagnostic testing (Ekwall & Shanker, 1976; Thomas, 1981).
   - expanding the domains of testing to include attitudes, feelings, impressions, etc. as they relate to the intended curriculum.
   - expanding response formats beyond the closed ended item, e.g., open-ended, short-answer and/or essay-type response (perhaps printed out for on-the-spot or later teacher analysis and feedback).

   The creative assessment of teaching and learning has been possible for some time; technology can make it more feasible and efficient. Why not make use of it, then, in more ways than just simple question-answer formats?

4. The process, of course, is not so simple. The basis for achieving a "working consensus" has been a matter of considerable philosophical debate in the arena of epistemology, i.e., what constitutes knowledge and the means whereby it is obtained. The position advocated here is multi-paradigmatic. It embraces both quantitative (traditional research using experimental and correlational designs and statistical analyses) and qualitative (naturalistic research using ethnographic, case study, and observational techniques and interpretation) methodologies, so long as a critical perspective is maintained. By this I mean a rigorous and sustained dialogue that addresses such questions as: What goes on in the name of curriculum? How did it come to be that way? Whose interests are being served by the way it is organized? Is this the way we want it to be? We have used the term critical inquiry to describe this multi-paradigmatic perspective. The interested reader is referred to Sirotnik and Oakes (1983).

5. Two projects particularly stand out in this regard: The Curriculum, Computers and Collaboration Project (see the "criteria-by-commonplace" matrix in THE PARTNERSHIP Newsletter, 1 (3), p. 7) and the curriculum inquiry task force of THE PARTNERSHIP.

6. This value/belief is directly traceable to Dewey (e.g., 1938) and the progressive education movement. See, also, the more recent extrapolation of these ideas in, for example, computer education and Papert's (1980) "microworlds" and mathematics curriculum inquiry and Romberg's (1983) "story shells."

References

Dewey, J. Experience and education. New York: Collier, 1983.

Edwards, L. Three steps to courseware selection. School Microcomputing Bulletin, 1984, 3, 185-186.

Ekwall, E. E., & Shanker, J. L. Diagnosis and remediation of the disabled reader. Boston: Allyn and Bacon, Inc., 1976.

Glaser, R. The future of testing: A research agenda for cognitive psychology and psychometrics. American Psychologist, 1981, 36, 923-936.

Goodlad, J. I. The dynamics of educational change. New York: McGraw-Hill, 1975.

Goodlad, J. I., & Associates. Curriculum inquiry: The study of curriculum practice. New York: McGraw-Hill, 1979.

Goodlad, J. I., Klein, M. F., & Tye, K. A. The domains of curriculum and their study. In J. I. Goodlad, et al., Curriculum inquire. New York: McGraw-Hill, 1979.

Goodlad, J. I. A place called school: Prospects for the future. McGraw-Hill. 1983.

Heckman, P. E., Oakes, J., & Sirotnik, K. A. Expanding the concepts of renewal and change. Educational Leadership, 1983, 40, 26-32.

Hively, W. Hively's choice. Elizabethtown, PA: Continental Press, 1983.

Klein, M. F., Tye, K. A., & Wright, J. E. A study of schooling curriculum. Phi Delta Kappan, 1979, 61, 244-248.

Marshall, W. K. Assessing educational software. Paper presented at the conference of the American Educational Research Association, 1984.

Mayer, R. E. Human cognition and the use of new technologies. CSE Report No. 232. Los Angeles: Center for the Study of Evaluation, UCLA, 1984.

Merrill, M. D. Don't bother me with instructional design: I'm busy programming. CSE Report No. 233. Los Angeles: Center for the Study of Evaluation, UCLA, 1984.

Oakes, J., & Schneider, M. Computers in the classroom: Another case of the more things change the more they stay the same. CSE Report No. 238. Los Angeles: Center for the Study of Evaluation, UCLA, 1984.

Oettinger, A. G. Run, computer, run: The mythology of educational innovation. Cambridge: Harvard University Press, 1969.

Papert, S. Mindstorms: Children, computers, and powerful ideas. New York: Basic Books, 1980.

Romberg, T. A. A common curriculum for mathematics. In G. D. Fenstermacher & J. I. Goodlad (Eds.) Individual differences and the commom curriculum. Chicago: University of Chicago Press, 1983.

Sarason, S. B. The culture of the school and the problem of change. Boston: Allyn and Bacon, 1971 (1st edition) and 1982 (revised edition).

Sirotnik, K. A., & Oakes, J. A contextual appraisal system for schools: Medicine or Madness? Educational Leadership, 1981, 39, 164-173.

Sirotnik, K. A., & Oakes, J. Critical inquiry and school renewal: A liberation of method within a critical theoretical perspective. Los Angeles: Laboratory in School and Community Education (Occasional Paper No. 4), UCLA, 1983.

Thomas, R. M. A model of diagnostic evaluation. In A. Lewy & D. Nevo (Eds.) Evaluation roles in education. London: Gordon & Breach, 1981.

Tyler, R. W. Basic principles of curriculum and instruction. Chicago: University of Chicago Press, 1949.

Van Buskirk, T. Hands-on beginner's manuals. T V Ontario, Ontario Educational Communications Authority, 1983.

Wilcox, R. R. Measuring mental abilities with latent state models. Journal of Mathematical and Management Sciences, 1984, in press.