INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University Microfilms International
300 N. Zeeb Road
Ann Arbor, MI 48106
INSTRUCTIONAL CHARACTERISTICS CRITICAL FOR PROMOTING ACHIEVEMENT IN COMPUTER COURSEWARE DESIGNED FOR GRADES 9-12

The Ohio State University

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106

Copyright 1985 by Sherman, Sallie Joyce

All Rights Reserved
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark □.

1. Glossy photographs or pages □
2. Colored illustrations, paper or print □
3. Photographs with dark background □
4. Illustrations are poor copy □
5. Pages with black marks, not original copy □
6. Print shows through as there is text on both sides of page □
7. Indistinct, broken or small print on several pages □
8. Print exceeds margin requirements □
9. Tightly bound copy with print lost in spine □
10. Computer printout pages with indistinct print □
11. Page(s) □ lacking when material received, and not available from school or author.
12. Page(s) □ seem to be missing in numbering only as text follows.
13. Two pages numbered □. Text follows.
14. Curling and wrinkled pages □
15. Dissertation contains pages with print at a slant, filmed as received □
16. Other □

________________________________________________________________________
________________________________________________________________________

University
Microfilms
International
INSTRUCTIONAL CHARACTERISTICS CRITICAL FOR PROMOTING
ACHIEVEMENT IN COMPUTER COURSEWARE DESIGNED FOR GRADES 9-12

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate School
of The Ohio State University

BY
Sallie Joyce Sherman, B.A., M.ED.

****

The Ohio State University
1985

Reading Committee
Keith A. Hall
Larry E. Miller
Victor M. Rentel

Approved By

Adviser
Educational: Theory and Practice
DEDICATION

This volume is dedicated to
Virginia Lewis Sherman and Polly Sherman Schiff,
for their constant support and encouragement,
and to
David and Blair Schiff,
so that they may know that some adventures are worth taking.
ACKNOWLEDGEMENTS

This work would not have been possible without the expert guidance of my outstanding committee. I wish to thank:

Keith A. Hall, my adviser and mentor, for providing excellent guidance and endless support;

Larry E. Miller for sharing his research expertise and wonderful stories; and,

Victor M. Rentel for insisting that a healthy intellect needs to take risks.
VITA

September 28, 1944 ............. Born - Aiken, South Carolina

1966 ............................. B.A. Centre College of Kentucky
Danville, KY (English, Psychology)

1966-1975 ........................ Teacher, Oak Hills High School
Oak Hills Local School District
Cincinnati, OH

1970 .............................. M.Ed. Xavier University
Cincinnati, OH (Counseling)

1975-1978 ........................ Post-graduate work
Xavier University, Cincinnati, OH
(Supervision, Administration)

1975-1982 ........................ Secondary Language Arts Supervisor
Oak Hills Local School District
Cincinnati, OH

1977-1979 ........................ Instructor, Cincinnati Technical
College, Cincinnati, OH

1983 .............................. Instructor, Summer Tech Program,
Columbus Public Schools,
Columbus, OH

1984 .............................. Director of Research and
Development, Mentor Technologies
Columbus, OH

1982-1985 ........................ Graduate Research Associate
Instructional Design and
Technology, The Ohio State
University, Columbus, OH

PUBLICATIONS


FIELDS OF STUDY

Major Field: Instructional Design and Technology

Organizational Implications of Technology
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEDICATION</td>
<td>i</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
</tbody>
</table>

## CHAPTER

**I. INTRODUCTION**

- Need for the Study .................................. 4
- Research Problem .................................. 5
- Research Questions .................................. 5
- Definition of Terms .................................. 6
- Limitations of the Study ............................ 7
- Basic Assumptions .................................. 8

**II. REVIEW OF THE LITERATURE**

- Attempts to Develop Evaluation Guidelines or Standards ........................................... 11
- SEDL/NEREX Evaluation Continuum ...................... 12
  - Level One Evaluation Efforts ...................... 14
    - MicroSIFT ........................................ 14
    - EPIE ........................................... 16
    - National Center for Vocational Education ... 20
      - Alberta, Canada ................................ 21
  - Level Two Efforts .................................. 22
  - Level Three Evaluation Efforts .................... 23
  - Level Four Evaluation Efforts ..................... 24
- Other Efforts to Develop Evaluation Guidelines ... 25
- Implications for Courseware from the Psychology and Instruction Literature ..................... 31
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Delphi Technique</td>
<td>32</td>
</tr>
<tr>
<td>Types of Delphi</td>
<td>33</td>
</tr>
<tr>
<td>Uses for the Delphi</td>
<td>34</td>
</tr>
<tr>
<td>Delphi Applications in Education</td>
<td>35</td>
</tr>
<tr>
<td>Delphi Process</td>
<td>35</td>
</tr>
<tr>
<td>III. METHODOLOGY</td>
<td>39</td>
</tr>
<tr>
<td>Panel Selection</td>
<td>39</td>
</tr>
<tr>
<td>Instrumentation</td>
<td>41</td>
</tr>
<tr>
<td>Round One</td>
<td>42</td>
</tr>
<tr>
<td>Round Two</td>
<td>44</td>
</tr>
<tr>
<td>Round Three</td>
<td>44</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>45</td>
</tr>
<tr>
<td>IV. FINDINGS</td>
<td>47</td>
</tr>
<tr>
<td>Procedures</td>
<td>47</td>
</tr>
<tr>
<td>Results</td>
<td>51</td>
</tr>
<tr>
<td>Refinement of Instructional Characteristics</td>
<td>56</td>
</tr>
<tr>
<td>Questions Requiring Recall, Analysis, Synthesis, and/or Evaluation</td>
<td>56</td>
</tr>
<tr>
<td>Opportunities to Imitate Models</td>
<td>57</td>
</tr>
<tr>
<td>On-line Option for Overview of Concepts to be Learned</td>
<td>57</td>
</tr>
<tr>
<td>Lessons Organized for Each Learner from the Least</td>
<td>57</td>
</tr>
<tr>
<td>Difficult to the Most Difficult Learning Task by Some Mechanism</td>
<td>58</td>
</tr>
<tr>
<td>On-line Overview of Skills to be Learned</td>
<td>59</td>
</tr>
<tr>
<td>Graphic and Visual Cues</td>
<td>59</td>
</tr>
<tr>
<td>On-line Statement of Goals and Objectives</td>
<td>60</td>
</tr>
<tr>
<td>Opportunities for the Teacher to Author Some Input</td>
<td>61</td>
</tr>
<tr>
<td>Organizers Embedded Within the Text</td>
<td>62</td>
</tr>
<tr>
<td>Portrays Main Point Visually</td>
<td>62</td>
</tr>
<tr>
<td>Verbal Imagery</td>
<td>63</td>
</tr>
<tr>
<td>Learner Control of Instructional Content</td>
<td>63</td>
</tr>
<tr>
<td>Presentation of Real-world, not Contrived</td>
<td>64</td>
</tr>
<tr>
<td>Information</td>
<td>65</td>
</tr>
<tr>
<td>Open-endedness</td>
<td>65</td>
</tr>
<tr>
<td>Use of Humor</td>
<td>65</td>
</tr>
<tr>
<td>Field-tested Materials which have Demonstrated Success with the Target Population</td>
<td>65</td>
</tr>
<tr>
<td>Agreement and Disagreement Between Groups</td>
<td>66</td>
</tr>
<tr>
<td>V. SUMMARY, IMPLICATIONS, AND RECOMMENDATIONS</td>
<td>72</td>
</tr>
<tr>
<td>Procedures</td>
<td>72</td>
</tr>
<tr>
<td>Discussion and Implications of Findings</td>
<td>75</td>
</tr>
<tr>
<td>Confirmation of the Research Literature</td>
<td>76</td>
</tr>
<tr>
<td>Importance of Defining Evaluation Criteria</td>
<td>79</td>
</tr>
<tr>
<td>The Need for a Clear Definition of the Target Audience</td>
<td>81</td>
</tr>
<tr>
<td>Consequences of the Debates Regarding the Purpose of Courseware and the Nature of the Medium</td>
<td>82</td>
</tr>
<tr>
<td>Implications of Individualized Instruction</td>
<td>83</td>
</tr>
<tr>
<td>Implications for Research</td>
<td>86</td>
</tr>
<tr>
<td>Recommendations</td>
<td>88</td>
</tr>
</tbody>
</table>

REFERENCES | 90 |

APPENIXES

A. Correspondence with Nominators of the Expert Panel | 94 |
B. Expert Selection Criteria and Nomination Ballot | 97 |
C. Synopsis Used in Nomination Mailing | 99 |
D. List of Expert Panel Members | 101 |
E. Correspondence with Expert Panel | 104 |
F. Pilot Test Instrument Review Panels | 109 |
G. Round One Delphi Instrument | 110 |
H. Round Two Delphi Instrument | 124 |
I. Round Three Delphi Instrument | 151 |
J. Round Three Summary | 185 |
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Panelists' Positions on Instructional Characteristics</td>
<td>53</td>
</tr>
<tr>
<td>2</td>
<td>Description of Agreement and Disagreement Among Groups</td>
<td>67</td>
</tr>
</tbody>
</table>
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SEDL/NEREX Continuum of Courseware Evaluation Efforts</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>Transformation of Certainty Scores</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>51</td>
</tr>
</tbody>
</table>
CHAPTER I

Introduction

Educators have been struggling with the process of evaluating and selecting instructional materials for years. In fact, most school districts have developed elaborate evaluation and selection procedures involving both educators and community members. Traditionally, these selection committees have evaluated prospective materials for (1) scholarship and (2) degree of congruence with curriculum and perceived community values; then submitted recommendations for adoption to the board of education.

Despite on-going efforts to clarify evaluation and selection procedures, the rapid influx of microcomputers into K-12 classrooms has raised new, difficult evaluation questions. First, who should be primarily responsible for evaluating and selecting courseware? Microcomputer courseware (programs that teach) often are considered supplemental, audio-visual materials, typically selected by librarians, resource teachers, or media specialists (Talmage, 1981).

Although most educational programs presently on the market are drill-and-practice mathematics lessons (Blum, 1981; Cohen, 1983a; Bialo and Erickson, 1984), newer programs will eventually become increasingly complex and may represent entire curricula. At what
point, therefore, do these supplemental materials become basal material and require a more complex evaluation process? Anticipating problems in selecting courseware, states such as Arizona are drafting legislation to place courseware evaluation under the legal umbrella of the state textbook approval process ("Arizona Set to Vote", 1983). Other states such as Texas and North Carolina have developed statewide evaluation processes as well.

The second question of how to evaluate the instructional quality of courseware has proven far more difficult. Kansky, Heck and Johnson (1981) believed that, because interactive computer technology is fundamentally different from print, courseware would require more precise evaluation guidelines. The effectiveness of a textbook they argued, ultimately depended upon the teacher using it; therefore, evaluation guidelines for textbooks could be somewhat flexible. Computer-based education, on the other hand, would not always be dependent upon the same type of teacher control. Increasingly elaborate diagnostic and management systems within the courseware itself may control a student’s path through the lesson. As a result, courseware evaluation processes would need to examine not only the content but possible student paths as well.

Unfortunately, the size of the courseware industry (an estimated $100 million a year investment by major publishers) and the abundance of poor courseware accentuates the evaluation problem. Although schools have purchased microcomputers at amazingly fast rates (Becker, 1983), the quality of software to support them is poor (Blum, 1981;
Olds, 1983; Roblyer, 1981b; Bialo & Erickson, 1984). "Unhappily the development and implementation of software packages has not been commensurate with the sales of the hardware. The production of 'educational' software is a fractured cottage industry dominated by enterprising programmers, not educators" (Blum, 1981, p.4). The emphasis upon programming rather than instruction has led some to estimate that as much as ninety percent of the "educational software currently on the market is not worth buying. Much of it consists of misguided efforts to translate textbook materials into software. In other cases, the software is poorly conceived, poorly designed, and difficult to use" (Olds, 1983, p. 3). Because instructional software represents such a substantial investment for education, Olds believes evaluation is a critical issue for educators.

In an attempt to help educators with this issue, numerous organizations have developed evaluation guidelines, checklists and services. The CONDUIT Clearinghouse at the University of Iowa, Northwest Regional Labs (MicroSIFT Project), National Council of Teachers of Mathematics (NCTM), and Educational Products Information Exchange Institute (EPIE) represent four of the earliest and most frequently cited of these services.

Despite these popular, early efforts, numerous new evaluation guidelines appear regularly. In fact, the courseware director for the Minnesota Educational Computing Consortium (MECC) recently estimated that there may be 200 courseware evaluation guidelines in use. Most of these instruments have been developed without using established
validity and reliability procedures and have resulted in the proliferation of evaluation checklists and commercial services with varying emphasis, little reliability, and lack of agreement regarding instructionally significant elements of courseware.

Need for the Study

Wade (1980) argued that "evaluation implies standards. Evaluating an instructional program implies knowing the characteristics of a 'good' program" (p. 32). Because there currently exist no agreed upon criteria or standards for courseware evaluation (Cohen, 1983; Roblyer, 1981b), it is difficult "to make good (courseware) choices in a new domain where there is little shared experience and no shared language" (Olds, 1983, p. 41).

The lack of agreed upon standards or criteria is particularly troubling considering that Educational Products Information Exchange Institute (EPIE) estimates the current educational software market at approximately 40 million dollars, and although schools presently have limited software budgets, that trend is rapidly changing. If the stakeholders in the courseware evaluation process (i.e., courseware evaluators, subject matter specialists, instructional developers, and teachers) can not agree upon criteria from which they can judge courseware and the quality of courseware does not improve significantly, the predicted tripling of the educational software market by 1990 and the estimated growth of school courseware budgets portend monumental problems for education. Educators will have
difficulty defending their courseware choices, students will be subjected to inferior instruction, school courseware libraries will be filled with inadequate programs, and American education will once again be accused of wasting taxes and of failing to provide quality education in this technological age.

On the other hand, if courseware evaluators, subject matter specialists, instructional designers, and teachers can agree upon criteria for judging courseware; those criteria can provide the common language and shared experiences to guide courseware development, evaluation and curricular integration and, thus, help assure quality education for our students.

Research Problem

The problem investigated was: Which instructional characteristics within courseware promote achievement according to courseware evaluators, subject matter specialists, instructional designers, and teachers?

Research Questions

The questions investigated included:

1. What do courseware evaluation experts, subject matter experts, instructional designers, and teachers believe are the instructional characteristics critical for promoting achievement in courseware designed for grades 9-12?
2. To what extent do courseware evaluation experts, subject matter experts, instructional designers, and teachers agree regarding which are the instructional characteristics critical for promoting achievement in courseware designed for grades 9-12?

3. To what extent do courseware evaluation experts, subject matter experts, instructional designers, and teachers disagree regarding which are the instructional characteristics critical for promoting achievement in courseware designed for grades 9-12?

**Definition of Terms**

*Courseware* was defined as computer programs that teach or instruct. The term is subsumed under the more general category of computer software which includes all types of computer programs.

*Courseware evaluation experts* were defined as those persons nominated by editors of the generic computer-based education journals and by directors of the software clearinghouses.

*Subject matter experts* were defined as those individuals nominated by the presidents of the professional organizations in the four academic disciplines in the secondary school core curriculum: English, social studies, science, and mathematics. The experts had both subject matter expertise and experience using computers.

*Instructional design experts* were defined as those persons nominated by instructional software publishers. Four publishers were randomly chosen from the Eastman Publishing Company list of the top
ten selling educational software packages (Software Merchandising, 1983).

Teachers were defined as those teachers with experience using computers and nominated by presidents of the National Education Association and the American Federation of Teachers.

Consensus was defined as 100% agreement or disagreement among panelists in a given round.

Instructional characteristics were those characteristics (1) that had been identified in the cognitive psychology literature as likely to promote achievement and (2) which could be manipulated by the instructional designer. Additional characteristics submitted by the panel in round one were also included.

Limitations of the Study

This study used the Delphi technique, a frequently used method to determine expert opinion, to answer the research question. This technique, believed to be an appropriate choice for the study, had several limitations which should be mentioned. First, although the Delphi technique was designed to achieve consensus, the panel actually reached a compromise (Mitroff & Turoff, 1975). Second, despite numerous investigative strategies, a clear understanding of why people believed as they did was difficult to achieve with mailed Delphi instruments. Third, the quality of the interpretations of panelists' comments rested solely on the investigator.
Broader limitations included (1) limiting the study to courseware designed for grades 9-12 and (2) choosing not to study the relationship between instructional strategy (e.g., drill, tutorial, simulation) and instructional characteristics.

**Basic Assumptions**

The purpose of this study was to reach consensus among panel members regarding the instructional characteristics critical for promoting achievement in courseware designed for grades 9-12. Only those instructional characteristics (1) that had been suggested by a review of the psychology literature as likely to promote achievement and (2) that could be manipulated by the instructional designer were included. Student characteristics (e.g., readiness, cognitive style, intrinsic learner motivation) and implementation concerns were omitted from this study. Other uses of computers in education (e.g., computer-managed instruction, word processing, or administrative uses) were also excluded.

Finally, all instructional characteristics suggested by this study must be subjected to further research with students in grades 9-12. Although this study did not include such research, future studies should be developed to compare expert opinion against student performance.
Assessments of courseware quality, especially that designed for the microcomputer, have been bleak (Blum, 1981; Cohen, 1983a & 1983c; Olds, 1983; Roblyer, 1981b; Bialo & Erickson, 1984; Thiessen, 1984). Much of the blame for the poor quality has been attributed to: (1) development methods which have emphasized programming skills rather than instructional quality (Blum, 1981; Roblyer, 1981a & 1983); (2) the failure of developers to understand fully and then exploit the interactive nature of the machine (Olds, 1983); (3) educational philosophies which differ on whether computer-assisted instruction or discovery learning is the most appropriate use of the machine (Roblyer, 1983); and (4) the lack of concern for the integration of the courseware into the classroom (Caldwell, 1983; Roblyer, 1983).

These four conditions have resulted in a number of courseware weaknesses documented by several studies conducted at Teachers College, Columbia University in conjunction with EPICE, the Educational Products Information Institute (Blum, 1981; Cohen, 1983a; Bialo & Erickson, 1984). Among the weaknesses cited in the most recent study were:
- a major emphasis on arithmetic programs designed for grades K-8,
- most packages were designed to provide supplementary drill and practice,
- most package objectives were directed at recall learning,
- few programs had been field tested during development
- there were not enough examples or demonstrations within the lessons,
- few programs had adequate suggestions for teachers,
- the user had little opportunity to get help, review directions, or control the presentation rate,
- few lessons had feedback designed to remediate,
- few lessons made effective use of audio, and
- few lessons had a management system or produced evaluation results which would be helpful for placing students within a curriculum sequence (Bialo & Erickson, 1984).

One must certainly question, however, the validity of the generalizations which resulted from the works of Bialo and Erickson (1984) and Cohen (1983a). For example, in the 1982 Exxon study (Cohen, 1983a), Cohen had suggested market trends by comparing the results of that study, which evaluated 29 programs, with the results of her 1981 study, which evaluated only six programs. The Bialo and Erickson study also suggested market trends but evaluated only 163 of the estimated 10,000 programs available. Conclusions from all these studies must be interpreted cautiously since none of the programs evaluated were randomly selected and, therefore, no generalizations to the total software market should have been made.

Nevertheless, EPIE has been the only organization which has studied and documented market trends for courseware quality. Because the works of Cohen and Bialo and Erickson were used as the basis of EPIE's highly publicized educational microcomputer software evaluation
service, EPIE Pro/FILES, their findings have had considerable impact on the courseware evaluation field.

Attempts to Develop Evaluation Guidelines or Standards

Wade (1980) noted "evaluation implies standards. Evaluating an instructional program implies knowing the characteristics of a 'good' program" (p. 32). Stufflebeam (1980) listed seven possible general benefits of standards. Those benefits included:

- a common language to facilitate communication and collaboration in evaluation,
- a set of general rules for dealing with a variety of evaluation problems,
- a conceptual framework by which to study evaluation,
- a set of working definitions to guide research and development of the evaluation process,
- a public statement of the state of the art of educational evaluation,
- a basis for accountability by evaluators, and
- an aid to developing public credibility for the educational evaluation field (p. 3).

Although numerous organizations have attempted to develop courseware evaluation guidelines, there are no agreed upon criteria or standards. Roblyer (1981b) contended that such standards have been more difficult to develop than expected and that existing criteria represent fundamentally different points of view about minimum requirements, the nature of the ultimate program, and the impact of technological change on future programs.

Cohen (1983a) noted, however, that the various guideline development attempts through 1982 represented a fundamental similarity and defined the new technology in terms of educational effectiveness.
Nevertheless, Caldwell (1983) maintained that most of the criteria used thus far to evaluate software were based in speculation and intuition rather than through research and evaluation. Despite these problems, the development of courseware evaluation guidelines continues (even Caldwell recently chaired a national committee to establish courseware evaluation guidelines for English teachers.)

Although only a few of the many evaluation efforts have been documented, the literature offers some help in understanding the various evaluation strategies and philosophies.

**SEDL/NEREX Evaluation Continuum**

A recent NIE project conducted by the Southwest Educational Development Laboratory (SEDL) and the Northeast Regional Exchange (NEREX) was particularly helpful in synthesizing courseware evaluation efforts. The *Evaluation of Educational Software: A Guide to Guides* (Jones & Vaughan, 1983), produced as a result of the joint project, used a continuum to help describe the degree of sophistication of the evaluation efforts. The continuum ranged from the most systematic efforts to the most flexible (Figure 1).
Level 1 Level 2 Level 3 Level 4
MicroSift EPIE School Microware Review MECC CONDUIT SOFTSWAP Professional Organizations & Publishers
Courseware Report Card SOFTSWAP

Systematic Flexible

Figure 1. SEDL/NEREX Continuum of Courseware Evaluation Efforts

Level one evaluation efforts were the most systematic evaluation efforts and included MicroSift (sponsored by the Northwest Regional Laboratory) and EPIE (the Educational Products Information Exchange). Both groups (1) systematically developed and field tested evaluation criteria, (2) trained reviewers at various sites across the country, and (3) published courseware reviews. Had recent systematic efforts by the National Center for Vocational Education (Chase, Gordon, and Makin, 1984) and the Alberta, Canada project (Thiessen, 1984) been completed earlier, they too would have been included in this group.

Level two evaluation efforts included such products as School Microware Review and Courseware Report Card, which contained evaluations performed by staff members and outside reviewers using a given set of criteria which had not been systematically developed and field tested. These publications were unique in that they encouraged courseware publishers to respond to the reviews in print.

Level three evaluation efforts consisted of organizations such as the Minnesota Educational Consortium (MECC), SOFTSWAP, and CONDUIT.
All these groups, whose primary purpose was to develop and distribute courseware, developed evaluation criteria to strengthen their own products. Because evaluation was not the primary concern of these organizations, the SEDL/NEREX project placed them further down the continuum.

Finally, the level four efforts or the professional organizations and publishers, such as the National Council of Teachers of Mathematics (NCTM) and Scholastic, developed evaluation forms "to be used by groups or individuals at their own discretion" (Jones & Vaughan, 1983, p. 11). Although these forms attempted to identify evaluation guidelines, no attempt was made to field test the instruments or train reviewers; thus, these organizations represented the least systematic end of the evaluation continuum.

In order to better understand the range of courseware evaluation efforts, one must examine each of the major efforts more thoroughly. The level one evaluation efforts were examined first.

Level One Evaluation Efforts

The MicroSIFT Project. The Microcomputer Software and Information for Teachers project (MicroSIFT) was begun in December 1979. According to Olds (1983), "hopes were high that this federally-funded software clearinghouse, housed at the Northwest Regional Laboratory, would bring organization and clarity to software evaluation. It was the first attempt of its kind." (p. 6) Through a lengthy process, MicroSIFT modified the original CONDUIT evaluation
form and process, then field tested the instrument at the MicroSIFT network evaluation sites across the country.

According to Holznagel (1983), director of the MicroSIFT Project, there have been approximately nine versions of the MicroSIFT guidelines. The 1983 Evaluator's Guide, which contained the evaluation criteria and an explanation of how to interpret them, was the most recent. The MicroSIFT evaluation form was divided into three sections: content characteristics (3 criteria), instructional characteristics (10 criteria), and technical characteristics (6 criteria). The criteria were rated on a Likert scale with additional space provided for assessments of potential for classroom use, strengths and weaknesses, estimates of time required, and general comments. Sections for a numerical evaluation of quality and an overall recommendation also were included.

The MicroSIFT courseware review process required a minimum of three or four months and included three stages. First, the MicroSIFT staff took a cursory look at the lessons. Those which passed that sifting stage were then more fully examined and described by the staff using the project Description Form. Finally, courseware was distributed to sites throughout the MicroSIFT network for review by volunteer teachers with experience in the subject and the grade level of the material. The computing expert at each site gathered the teacher reviews and then wrote a summary for the MicroSIFT evaluation of the package. Generally, students were not involved in these first
three stages of evaluation. A more in-depth fourth stage involving students was planned.

Olds' assessment of the MicroSIFT Project was worth noting:

MicroSIFT, as the first attempt of its kind broke new ground and illuminated problems yet to be solved: first, evaluation by committee may be a worthwhile process if the committee can hammer out a consensus. But if the committee never resolves differences and the consensus is created by another party, (the computer expert's summary review) the result may be a very odd stew. Second, while the goal of MicroSIFT'S methodology — to attain reliability and credibility by reducing the level of subjectivity in evaluation — is a worthy one, the result does not always make clear distinctions among different programs or differentiate good programs from bad ones. In spite of its wrinkles, MicroSIFT has provided educators with one of the first evaluation forms to attempt to collect data in some depth. In addition, the creation of a network of educators interested in evaluating software and sharing their evaluations was valuable in itself, and contributed to the growing body of knowledge about evaluation as a process. (Olds, 1983, pp. 6-7)

EPIE. Another major courseware evaluation project was begun by EPIE in the early 1980's. EPIE, since 1967, had provided analyses of textbooks and other instructional materials for elementary and secondary schools. During 1981, the EPIE Institute, in a joint project with Teachers College, produced its first microcomputer courseware evaluations based on an instrument adapted from EPIE'S original materials analysis instrument. The revised instrument, developed by Cohen as part of her dissertation research, asked analysts to review courseware on approximately 80 criteria in the following areas:
General categories for a summary and recommendations were also included.

The instrument was field tested with six microcomputer packages by seven computer analysts, equally divided between those who had a computing/mathematics background and technical programming expertise and those who had extensive experience in instructional materials development and analysis (Blum, 1981). After each lesson had been reviewed, the analysts discussed their evaluations. The first analysis reflected a 72% inter-rater reliability; the sixth, a 96% reliability. The results of these reviews were summarized and published in Pro/FILES. Meanwhile, EPIE joined with Consumer's Union and a consortium of school districts in Albuquerque, Boston, Cincinnati, Houston, Detroit, and Salt Lake City to evaluate both
hardware and software products and began training teams in these districts to develop courseware analysts.

In 1982 Cohen extended her original study at the Microcomputer Resource Center at Teachers College in a larger project funded by EXXON (Cohen, 1983a). This study differed from the first in several ways. First, the EXXON project examined 29 rather than 6 programs marketed for school use. Second, the evaluation teams were expanded to include a subject matter expert, instructional designer or educational psychologist, a practicing teacher, and Cohen, the project director.

Cohen saw the addition of the practicing teacher as particularly significant. Especially valuable were the teacher comments regarding teacher training, curricular integration, ease of implementation and pedagogical concerns. Cohen trained the teachers herself and served on the evaluation teams to increase reliability.

As the Project Director, I would look at each program long enough to become familiar with it and I would jot down detailed notes about its benefits and its flaws. I maintained the consistency throughout each team and contributed a knowledgeable background and guiding force throughout the project. (Cohen, 1983a, p.6)

The third way the EXXON study differed from her earlier one was the use of students to verify or refute the evaluators' findings. People on the teams were asked to provide students of the age and characteristics intended for the program. These students, in addition to the students in the Teachers College special programs created a pool of student volunteers. Each package reviewed was used by a
student and data were collected. Student evaluations lasted one-half to one hour each. As Cohen stated, this phase was the:

most fun, the most interesting, and the most informative. We always tried to validate our assumptions by purposely choosing a particular type of student or specific part of a program to try out. Therefore, the evaluators would usually have definite question(s) in mind before the actual tryout (p. 7).

The fact that students were purposely chosen to validate the assumptions of the researchers however, raises serious questions about the validity of the student data and the possibilities for researcher bias.

Nevertheless, following an evaluation debriefing, Cohen compiled the final report, noted whether or not consensus was achieved, and submitted the evaluation to the EPIE director, who "transformed it into final copy." Evaluators were allowed to correct or edit the director's final report. Thus, the same problems that Olds suggested were inherent in MicroSIFT's third party synthesis of the evaluations were present in the EPIE process as well.

Cohen's findings made assessments of courseware quality and market trends, despite lack of random sampling and the validity problems discussed earlier. Yet, she believed her work "helped to establish guidelines and standards for the new field" (p. 4) and EPIE has continued to use Cohen's original work as the basis of the EPIE evaluation process. Despite Cohen's claim of developing new standards for the field, Olds found the EPIE evaluations disappointing.

At this time, I have only one sample of their software evaluations to consider. Unfortunately, from this one
sample, the standards that were set by EPIE'S earlier efforts are not met in their new format or through their new methodology. The comments are briefer and made to sound more objective. A quantitative summary rating of major attributes is provided. There are brief quotes from magazine reviews and from student users, and there is one page of brief summary statements on various attributes of the program. While the format appears more accessible and there is less reading required, the overall impression is now more like the MicroSIFT evaluations, that is, ambiguous and uncertain. And, sad to say, the reader does not know after reading the evaluation whether the program is worth serious consideration. (p. 7)

**National Center for Vocational Education.** Because most of the evaluation processes had concentrated on courseware designed for the core curriculum areas, the National Center for Vocational Education sought and received funding from the United States Office of Vocational and Adult Education to develop a system for evaluating microcomputer courseware for vocational and technical education. The project (1) reviewed the courseware evaluation literature, (2) examined existing evaluation guidelines, (3) visited major courseware evaluation sites, (4) developed a three-part evaluation form which was refined by two panels of consultants representing courseware evaluation and vocational and technical education, (5) pilot tested the instrument, and (6) suggested a procedure for using the instrument (Chase et al, 1984).

Although the purpose of the project was to evaluate courseware, the process also included evaluation of applications software critical for vocational and technical education such as word processing, file managers, and spreadsheets. This project represented one of the most systematic of its kind and the researchers sought to garner crucial
political support by including stakeholders and facilitators within vocational and adult education throughout the project.

Alberta, Canada. Another recent project designed to systematically evaluate courseware was part of the Computer Technology Project sponsored by the Canadian Alberta Education Department (Thiessen, 1984). Although the department considered courseware development a high priority, the project surveyed existing courseware developed for the Apple II family of computers, the provincial standard, in order to avoid duplication of efforts and to provide teachers data from which they could make informed purchase decisions.

The project team began evaluating programs in K-12 mathematics and eventually extended into other curricular areas. The three stage evaluation process adapted the MicroSIFT model but used the final stage of the courseware evaluation process to evaluate for congruency with the Alberta curriculum. Teams of teachers were trained to use the evaluation guidelines and then were paid for their evaluations. Those programs which received ratings of "prescribed" or "recommended" were purchased in bulk and sold at a discount to the provincial schools. Formal evaluation reports were sent to the schools periodically.

By the end of March 1984, eight hundred packages had been evaluated and approximately 90% had been rejected. The reasons for the high rate of rejection included:

- inappropriate content range,
- poorly developed content sequence,
- superficial content and insufficient practice,
- inadequate student interaction,
- lack of feedback which remediated incorrect responses.  
(p. 5)

Although highly criticized by some participants at the 1984 National Educational Computer Consortium for being too systematic and blacklisting too many programs, the Alberta project represents one of the most systematic and professionally executed evaluation efforts to date.

**Level Two Evaluation Efforts**

Publications such as *Courseware Report Card* and *School Microware Review* represented a slightly different approach to courseware evaluation. These publications, also devoted exclusively to courseware reviews, relied upon reviews based upon specific criteria, although they did not train reviewers to use those criteria. Publishers were encouraged to respond comments by reviewers.

*Courseware Report Card* claimed to be "the first publication to offer a large volume of detailed critical reviews of educational software for a variety of microcomputer systems" (Jones & Vaughan, 1983, p. 329). The publication sought to evaluate a broader range of educational software than the evaluation efforts discussed thus far, including games, simulations, and authoring languages. In fact, the editors, feeling that drill and practice programs were not the most appropriate use of the computer, allowed reviewers to give no mark higher than a "C" on the appropriateness criteria for that type of program.
The evaluation format included: a short introduction, a thorough description, an evaluation on six criteria (performance, ease of use, error handling, appropriateness, documentation, and educational value), a box of basic program information, and a box for a letter grade (A-F) on each of the six criteria. All the editors/reviewers had classroom experience and a wide range of materials development and production experience. Olds (1983) considered this publication to be one of the most promising recent efforts to provide software evaluation.

Level Three Evaluation Efforts

Further along the SEDL/NEREX evaluation continuum were such groups as CONDUIT, SOFTSWAP, and MECC. These groups, primarily concerned with their own courseware development, distribution, and teacher training, developed evaluation criteria primarily to guide and evaluate their own product development. For example, CONDUIT, begun in 1971 with National Science Foundation support and affiliated with the University of Iowa, encouraged university faculty to develop courseware. The resulting programs then underwent a process of peer review for conceptual validity and instructional usefulness. Additionally, they were "technically reviewed for program accuracy and transferability to a wide range of computer systems" (Jones & Vaughan, 1983, p. 46). Recently, some of the CONDUIT courseware has been aimed at the senior high school market as well. Likewise SOFTSWAP and MECC,
both aimed at the K-12 audience, developed evaluation guidelines to improve their courseware development.

Level Four Evaluation Efforts

Numerous professional organizations and publishers developed courseware evaluation criteria and standards to meet the unique needs of their constituents. The National Council of Teachers of Mathematics developed one of the first and most popular of these guidelines. The authors stated that the NCTM Guidelines were derived from "state-of-the-art knowledge, the experience of users and producers of software, and a good dose of intuition" (Kansky, Heck & Johnson, 1981, p. 603). Their criteria included: instructional range, instructional grouping, execution time, program uses, user orientation (instructor and student), content, motivation and instructional style, and social characteristics. Yet, they believed courseware evaluation should be a matter of professional judgment in which criteria were used as a means to a decision rather than an end. By 1985 most other major academic associations had produced evaluation guidelines to be used by teachers in their respective disciplines.

Although not described in the 1983 NIE project, the National Education Association believed poor quality education courseware could be improved by adhering to "guidelines that incorporate the components of educational software design with the characteristics of computer games and the standards of technical quality" (Gold, 1984, p. 43). Thus, the organization published a guide for developing and
evaluating courseware in 1983. Their standards were divided into the following categories:

Technical Standards
  Software
  Technical Manual
Education Standards
  Time Requirements
  Bias-free Content
  Technical Instructional Features
Documentation-Teacher's Manual
  Teacher Instructions
Software Design

The literature reported no systematic attempt to validate evaluation criteria or train teachers in using the NEA criteria.

Other Efforts to Develop Evaluation Guidelines

Of the various evaluation efforts discussed thus far, most emphasized the judgments of the reviewers and a few even included student data. Della-Piana and Della-Piana (1982) at Northwest Regional Laboratory dramatically shifted the prevailing research focus by concentrating on gathering student data and "making courseware transparent" so the prospective user could make the most informed final choice. Believing that the criticism which "helps us to see how things work" is most needed, they stated:

the function of criticism is to show how things work, to make them transparent, to demistify. But it is not enough to show how things work. The portrayal of the workings of a piece of courseware must have utility for two audiences. It must involve users in making their own interpretations, judgments and adaptations of courseware for their own purposes. It must involve developers and distributors in "revision," in seeing the
work again or looking back at it from a perspective that may drive redevelopment (p. 2).

Thus, they contended the best way to study courseware was to trace student paths by reviewing hard copy records of those paths and by flowcharting the narrative. As a result, the Della-Pianas gathered data sets in three areas: what (1) was designed to happen (author-defined domain), (2) did happen (student-defined domain), and (3) should have happened (ideal-defined domain). Relying upon six assumptions, they proceeded to gather unique data for the field of courseware evaluation. In addition, the assumptions underlying their study suggested new directions for future courseware evaluation. They assumed that:

1. courseware users have limited information processing capacities,
2. there as yet is no prescriptive science of design, thus, implications derived from cognitive, developmental or behavioral psychology are still only hypotheses which have yet to be verified,
3. the most valuable information about courseware may be the learner's path through the lesson,
4. courseware standards must be derived from many applications in many different subject areas,
5. courseware structure for any given piece of software will vary for different audiences,
6. typically courseware evaluation will require naturalistic inquiry and include an examination of the curricular integration, characteristics of the learner, relationship of worth and cost, and generalizability across different contexts. (pp. 4-5)

Thus, the Della-Pianas shifted the focus of courseware evaluation from the traditional methods or evaluation, which examined the instructional design of a lesson, to a new method which examined the author, student and ideal domains of a lesson.
The Della-Pianas believed, as do many others, that the rapidly changing hardware technology and software design would make any existing courseware criteria obsolete. In order to help users decide which package to use, however, the researchers believed it helpful to create dissonance by juxtaposing the author, student, and ideal data sets after the courseware had been screened and found to be "educationally valuable, technically competent, thorough in documentation, and widely used" (p. 28). They did not explain how these latter criteria should be determined. For those finding the in-depth analysis too time consuming, the Della-Pianas developed a "Microcomputer Courseware Evaluation Open Checklist." This checklist, which required training for use and examples of its use in many different settings, was designed to analyze only the student's path through the lesson. As a result, the Della-Piana study suggested a change in the direction of courseware evaluation.

Wade (1980) and Jay (1983) both argued for criteria which emphasized the findings of cognitive psychology research. Wade argued specifically for using nine instructional events described by Gagne and Briggs as a framework for evaluation:

1. gaining attention,
2. informing the learner of objectives,
3. stimulating recall of prerequisite learnings,
4. presenting the stimulus material,
5. providing learning guidance,
6. eliciting performance,
7. providing feedback about performance correctness,
8. assessing the performance, and
9. enhancing retention and transfer.
Jay advocated five principles and supporting practices from cognitive psychology research as criteria for courseware evaluation. The five principles included:

1. Short term memory is limited by space and time,
2. Student performance is a function of the type of message, the intent of the message, and the type of student,
3. Graphics enhance memory,
4. Processing abilities mature with age, and
5. Information about performance and how to correct and incorrect responses is helpful. (Jay, 1983, pp. 23, 24, 25)

Roblyer (1981b) argued for a more practical approach to courseware evaluation guidelines. She maintained reviewers should be using (1) instructional criteria which should be applied to any courseware, regardless of purpose, and (2) differential criteria which allow for considerable variation in presentation and appearance, depending on the purpose of the courseware (p. 53).

Thus, the instructional setting and the target population should influence the evaluation criteria used.

Rather than list guidelines, Roblyer discussed instructional concerns and minimum characteristics which should be considered in three categories of courseware criteria:

**Essential characteristics**
- Statement of objectives
- Statement of entry skills
- Design of learning activities
- Design of tests
- Content integrity
- Design presentation

**Aesthetic Characteristics**
- Spacing and format
- Use of color

**Differential characteristics**
- Learning events to include on-line
- Amount of learner control
- Type and extent of feedback
Response format
Amount of text on screen
Use of peripherals (multi-media)
Graphics and animation (p. 50).

Roblyer believed "the essential criterion is that students can use the resulting materials easily, can accomplish the objectives specified for the package, and enjoy learning with the materials" (Roblyer, 1983, p. 29).

Cohen (1983b), drawing upon her work with EPIE and Roblyer's idea of differential criteria, developed a list of necessary attributes to consider in designing courseware for the microcomputer. Her list identified attributes generic to all instructional materials as well as those specific to courseware.

Finally, Olds (1983) in developing the overview for the SEDL/NEREX project, suggested his own list of the features of quality software:

1. It should be the outgrowth of a fully conceived and carefully articulated 'intellectual model of the content domain' (e.g., good science software should represent the best current scientific understanding),
2. It should reflect an understanding of the cognitive and developmental needs and capacities of the learner (e.g., younger children need more concrete representational experiences),
3. Since the computer provides an environment for interaction, its best pedagogical use should support the inherently interactive nature of knowledge construction for the learner (e.g., the computer is a poor medium for giving a lecture), and
4. It should make use of the special qualities of computer technology in truly functional ways (e.g., use graphics to make an abstract concept concrete). It should not attempt to carry out instructional tasks far better suited to other media (e.g., it should not 'be a textbook) (p. 3).
In conclusion, having reviewed the wide variation in approaches and criteria for courseware evaluation—from checklists designed by professional organizations for individual teacher use to elaborate, controlled evaluation projects such as MicroSIFT and EPIE—two conclusions seem justified. First, the vast majority of guidelines emphasized the technical rather than the instructional quality of courseware. Some have suggested this was because educators already knew how to evaluate instructional quality and needed more help with the technical aspects of the new technology. Whether or not that assumption was correct, technical standards are generally less difficult to develop than instructional ones. Second, a review of the literature suggested that, in light of the booming educational software market, there is a growing and immediate need to reach a consensus on those instructional characteristics most likely to influence courseware quality. Until that has been accomplished and until those characteristics have been subjected to research with students, additional courseware evaluation guidelines seem counterproductive. Glass (1978) may have been correct when he asserted, "Perhaps the only criterion that is safe and convincing in education is change. Increases in cognitive performance are generally regarded as good, decreases as bad" (p. 259); nonetheless, the question remains: What are the characteristics of courseware that bring about such change?
Implications for Courseware from the Psychology and Instruction Literature

Glaser (1976) reported several instructional conditions discussed in the psychological literature that promote student competence and have been debated in this Delphi study:

1. structured knowledge through the organization of instructional sequences,
2. problem solving activities,
3. reinforcement, and
4. instructional implementation which facilitate transfer of learning and mastery learning.

These conditions resulted in the inclusion of the Delphi items relating to the structuring of the content, opportunities for problem solving, corrective and immediate feedback, knowledge of goals and objectives for the lesson.

The Weil and Murphy (1982) review of instructional characteristics and implications of research regarding information processing, cognitive developmental, and behavioral strategies suggested additional characteristics for consideration. Those offering the greatest potential for instruction and courseware included (1) information processing strategies utilizing advance organizers; (2) strategies relating to how one organizes and attains information about concepts; (3) cognitive developmental strategies which allow learners opportunities to control and manipulate the learning environment, place the learner at the center of educational process, and promote individualized instruction; and (4) behavioral strategies related to modeling and reinforcement.
Additional findings reviewed by Weil and Murphy and reported in the teaching/effectiveness research literature suggested issues surrounding academic focus (e.g., de-emphasizing non-academically oriented materials such as games or puzzles), the need for teacher direction and control, concern for academic progress, and student accountability and cooperation. Thus, Delphi items graphic and visual cues, organizers embedded in the text, opportunities to ask questions of the program, overview of skills and concepts to be learned, learner control issues, use of humor, and opportunities for practice and success were included in the study.

Ripple and Drinkwater (1982) identified issues in the transfer of learning which also were evaluated in this study. Implications of this work resulted in the inclusion of such instructional characteristics as opportunities to elaborate and verbal imagery.

Thus, the round one Delphi instrument incorporated those instructional characteristics believed to promote achievement.

The Delphi Technique

The Delphi Technique, developed for an Air Force RAND Corporation study in the early 1950's, was designed to elicit and refine group judgments (Dalkey, 1969). Originally, the technique was developed to help the military determine the consequences of atomic attack on the United States. Rapid technological changes in the aerospace and technology industries required the military to develop a tool to facilitate the critical skill of forecasting. In the early 1960's,
non-military interest in the technique spread. The Delphi, seen by Linstone and Turoff (1975) as an art more than a science, was systematically tested by the RAND Corporation against other types of group decision making processes. Those studies found face-to-face discussions less accurate than individual estimates without discussion in such processes as the Delphi. Numerous other studies have supported the superior quality of the Delphi technique (Crawford & Cossitt, 1980).

Psychologists found numerous problems inherent in face-to-face discussions. These included: (1) the influence of dominant individuals, (2) semantic noise, and (3) group pressure for conformity (Dalkey, 1969, p. 14). The Delphi attempted to control for these problems by: (1) requiring anonymous response to limit dominant individuals, (2) controlling feedback to reduce semantic noise, and (3) statistically evaluating group response to reduce pressure for conformity and ensure individual participation.

Types of Delphi Studies

Typically, there have been two types of Delphi panels: the conventional panel, which relied upon analysis of mailed questionnaires, and the real-time panel, which used a computer. Whichever type used, the process generally involved four phases:

1. exploration of the subject and the addition of information by panelists,
2. attempts to gain insight into the amount of agreement and disagreement on the topic within the group by the researcher,
3. exploration of disagreement for possible reasons,
4. analysis and evaluation of information returned to
the panelists for reconsideration (Linstone and Turoff, 1975, pp. 5-6).

**Uses for the Delphi**

Linstone and Turoff (1975) listed several reasons for using the Delphi which helped determine the use of the technique for this study. First, the problem of courseware quality lended itself less to precise analytical techniques and more to subjective judgments. Second, because courseware evaluation is a complex problem requiring expertise from individuals of diverse backgrounds who have little history of communication, a method was needed to facilitate communication. Third, face-to-face meetings were considered too costly and time-consuming. Fourth, because courseware evaluation has raised philosophical and political issues about which people feel strongly, some method was needed to promote maximum communication by reducing peer pressure (p. 4). According to these authors, the process may lead to a sort of collective group intelligence.

Although the primary use of the Delphi has been for forecasting, a variety of other uses have proven helpful as well. Some of these have included:

- gathering current and historical data not accurately known or available,
- examining the significance of historical events,
- putting together the structure of a model,
- delineating the pros and cons associated with potential policy options, and
- exposing priorities of personal values, social goals (Linstone and Turoff, p. 4).
Delphi Applications in Education

Weaver (1971), in his survey of the use of The Delphi in education, believed the three best educational uses included:

1. a method for studying the process of thinking about the future,
2. a pedagogical tool or teaching tool which forces people to think about the future in a more complex way than they ordinarily would, and
3. a planning tool which may aid in probing priorities held by members and constituencies of an organization.

Believing that we must look at both the Delphi process and the product if we are to gain a thorough insight into a problem, Weaver contended that the technique may prove helpful because the controlled panel judgments would separate dreams from realities in the process. Because the participants would be experts, they would take into account new or discrepant information and construct logically sound deductions about the future based upon a thorough and disciplined understanding of particular phenomena and how they relate.

Weaver chronicled Delphi studies conducted within education prior to 1971 and showed that those studies reflected a wide range of purpose. Since that time, numerous other Delphi studies have extended the usefulness of the technique within education, especially for administrative decision-making.

The Delphi Process

Hostrop (1975) listed eight steps in a typical Delphi study:
1. identify the panel of experts,  
2. determine their willingness to participate,  
3. gather input,  
4. analyze data,  
5. mail the results of the input to panel and seek their reaction,  
6. analyze new input using descriptive statistical techniques,  
7. have group re-examine the analysis of the data, and  
8. analyze the new data (p. 68-9).  

Brooks (1979) believed steps two and three could be done simultaneously, and that three mailings would generally be enough to complete the study. He also suggested using three persons to analyze the data in order to avoid bias.

Numerous other suggestions for a successful Delphi study also appeared in the literature. Those incorporated in this study included:

1. keeping the idea generation and decision making processes separate (Crawford & Cossitt, 1980),  
2. creating panels which include "stakeholders: those who are or will be directly affected; experts, those who have an applicable speciality or relevant experience; and facilitators, those who have skills in clarifying, organizing, synthesizing, stimulating... plus, when it seems appropriate, individuals who can supply alternative global views of the culture and society" (Scheele, 1975, p. 68),  
3. designing ways of motivating the participants (Linstone & Turoff, 1975; Scheele, 1975) through tokens and the expert status of the panel,  
4. helping panelists create a group sense by providing meaningful responses which incorporated their ideas,  
5. keeping panelists focused upon key concepts through notes from the researchers,  
6. introducing ambiguities purposely to help refine thinking and seeking clarification for unexplained responses,  
7. using colored Delphi instruments,  
8. giving examples of responses,  
9. highlighting both agreement and divergence by graphically depicting the percent of agreement and disagreement on an item and by listing panelists' comments,
10. explaining the way the responses were being categorized,
11. praising or highlighting at least one response from each 
    participant to help them feel appreciated (Scheele, 
    pp. 68-71),
12. making the respondents feel their opinions were valid 
    by explaining how they were selected, reporting 
    their opinions, and commenting upon the quality of their 
    comments in the cover letters (Cyphert & Gant, 1970).

Despite the relative strengths of the Delphi Technique, use of 
the technique has presented numerous problems for some researchers. 
Brooks (1979) believed the most common problems centered around the 
capabilities of the panel and the time frame required. He contended 
the mailings required at least four to five months and that 
researchers sometimes did better allowing seven or eight months for 
the process. Linstone and Turoff believed researchers needed to 
carefully consider problems before beginning the technique, believing 
that perhaps as many Delphi studies have succeeded as have failed to 
reach consensus. Possible reasons for such failure might be:

- imposing monitor views and preconceptions of a problem 
  upon the respondent group by overspecifying the structure 
  of the Delphi and not allowing for the contribution of 
  other perspectives related to the problem,
- assuming that Delphi can be a surrogate for all other 
  communication in a given situation,
- poor techniques of summarizing and presenting the group 
  response and ensuring common interpretations in a given 
  situation,
- ignoring and not exploring disagreement so that 
  discouraged dissenters drop out and an artificial 
  consensus is generated,
- underestimating the demanding nature of a Delphi 
  and the fact that the respondents should be recognized 
  as consultant and properly compensated for their time if 
  the Delphi is not an integral part of their job function.
These authors believed it impossible to rid the Delphi of all problems, so the researcher must work to modify them as much as possible.

Finally, Weaver reminded researchers that Delphi studies project judgments not facts about the future. Clarifying and sharing assumptions, therefore, must become a critical part of the Delphi technique. Thus, like works of art, judgments of the future can be accepted or rejected but not proved false.
CHAPTER III
Methodology

This study used a modified three-round Delphi technique to answer the research questions. The growing body of cognitive research literature related to instructional characteristics likely to promote achievement allowed the researcher to develop a structured rather than an open-ended Delphi instrument.

Panel Selection

On April 20, 1984, nominations of experts to serve on the Delphi panel were sought from the leaders of four groups considered stakeholders in the courseware evaluation process: courseware evaluators, content specialists, instructional designers, and classroom teachers. The Classroom Computer News Directory of Educational Computing Resources (1983) was the source for the organizations within the nomination categories.

Nominators were sent (1) a brief letter explaining the nature and purpose of the study (Appendix A), (2) a synopsis of the study (Appendix C), (3) a nomination form (Appendix B), (4) a stamped return envelope, and (5) a felt-tip pen, which served as a token to thank nominators for participating. The nomination form listed the criteria
(Sutphin, 1981) for the selection of experts. Two nominations were sought from each of the nominators. Follow-up phone calls were made to those nominators who had not responded by the May 4 deadline.

Nominations for panelists were obtained in the following manner. First, nominations of experts in courseware evaluation were sought from both researchers and practitioners. Nominations from groups representing research were sought from the three professional associations within the United States (1) not linked with a particular academic group, (2) dedicated to educational computing, and (3) which published related professional journals. The nominators included the journal editors from the Association for the Development of Computer-Based Instructional Systems (ADCIS), the Association for Educational Data Systems (AEDS), and the Society for Applied Learning Technology. Nominations from groups representing courseware evaluation practitioners were sought from the directors of the four software clearinghouses: CONDUIT, Microcomputer Education Applications Network (MEAN), Microcomputer Software and Information for Teachers (MicroSIFT), and SOFTSWAP.

Second, nominations of content matter experts were requested from the presidents of the professional associations representing the four core curriculum areas: the National Council of Teachers of English, the National Council of Teachers of Mathematics, the National Council of Teachers of Science, and the National Council of Teachers of Social Studies.
Third, nominations of instructional design experts were sought from four publishers randomly chosen from the list of top ten best-selling instructional software packages (Software Merchandising, 1983). The four publishers randomly chosen included: Harcourt, Brace and Jovanovich, Inc.; Spinnaker Software Corporation; Developmental Learning Materials (DLM); and Edu-Ware Services, Inc.

Fourth, nominations for the experts representing classroom teachers were sought from the presidents of the American Federation of Teachers and the National Education Association.

From the twenty-eight persons who were nominated as experts, 18 were selected to serve on the expert panel through the following method: those receiving more than one nomination were automatically selected (1 member); those nominators nominating themselves were automatically selected (6 members); the remaining members were chosen by a disinterested third party knowledgeable in the field (11 members). The third party attempted to achieve a balanced representation of the four professional areas contained in the study. Of the original 18 panel members chosen, 13 participated (Appendix D).

On December 26, 1984, nominators were sent a letter thanking them for their nomination and notifying them of the beginning of the study (Appendix A).

Instrumentation

The Delphi instruments were designed according to methods prescribed by Dillman's total design method (Dillman, 1978) and were
revised as a result of a formative evaluation prior to each mailing. (See Appendix F for names of the formative review panels). Each mailing contained a cover letter, stamped return envelope, Delphi instrument, and token. Tokens for each round included, in the following order, an adhesive diskette holder; a $.50 McDonald's coupon; and a candy bar. The instruments were a different color for each round and the deadlines for returning the instruments were approximately one week after the anticipated date of delivery.

Round One

Round one, mailed January 1, 1985, consisted of two parts. Part one asked panelists to react to those instructional characteristics (1) that had been suggested by a review of the cognitive psychology literature as likely to promote achievement and (2) that could be manipulated by the instructional designer. Student conditions and learning outcomes (Reigeluth and Merrill, 1979) were omitted from the study. Also, panelists were asked to suggest additional instructional characteristics promoting achievement which they thought critical. Part two sought minimal information: work and home phone numbers, verification of area of expertise, and verification of address. All instruments were coded to provide the anonymity required by the Delphi technique. (See Appendix G for the round one instrument.)

In round one, broad instructional characteristics were used to allow panel input in the refinement process. The original list of
instructional characteristics suggested by the literature as likely to promote achievement included:

1. statement of goals and objectives,
2. increased progression of skills or concepts to be learned
3. use of humor,
4. depth or level of questioning,
5. overview of skills or concepts to be learned,
6. opportunities within the program to ask questions of the program,
7. organizers embedded within the text (e.g., questions, headings, sub-headings),
8. graphic or visual cues (e.g., color coding, diagrams, charts),
9. verbal imagery (e.g., metaphors, similes, analogies, descriptive phrases),
10. opportunities to imitate models,
11. knowledge of results,
12. corrective feedback (i.e., explanation of incorrect responses),
13. opportunities for practice,
14. high degree of learner control,
15. opportunities for problem solving,
16. opportunities to elaborate (i.e., further describe or build upon ideas within the lesson), and
17. opportunities for success.

For each instructional characteristic, panelists were asked (1) to agree or disagree with alternative positions, (2) to indicate their degree of certainty, (3) to support their position if they so desired and (4) to suggest additional characteristics.

Follow-up phone calls were made to obtain those instruments which had not been returned by Wednesday, January 9, the due date. Thirteen instruments had been returned by the round one cut-off date, approximately two weeks after the due date. Of the five missing instruments, one was returned to the sender with no forwarding address, two were promised to be returned but never were, one was
returned with a note explaining that she could not participate because she had become a full-time doctoral student, and one panelist had to withdraw because of a newly instituted company research policy.

Round Two

Round two was mailed February 9, 1985. In this round, panelists were (1) given the percentage of agreement and disagreement for each item in the previous round, (2) shown their positions from the previous round, (3) shown the comments from the previous round, (4) given the additional items suggested by the panel, and (5) given notes from the researcher in cases where clarifications, rewordings, or summations were needed. Panelists were then asked to reconsider their earlier positions and react to the new items suggested by their colleagues (Appendix H).

Follow-up postcards were mailed one week after the initial mailing, thanking those who had returned their instruments and reminding those who had not to do so. As predicted by Dillman (1978), the follow-up mailing generated the same amount of response as the initial mailing. Follow-up phone calls were made to the two panelists who still had not replied by March 1. All instruments were returned by March 6.

Round Three

In round three, mailed March 10, panelists were (1) given the percentage of agreement and disagreement for each item in the previous
round, (2) shown their positions from the previous round, (3) shown the comments from the previous round, and (4) given notes from the researcher in cases where clarifications, rewordings, or summations were needed. Panelists were then asked to reconsider their earlier positions. Items which had 100% agreement throughout two rounds were withdrawn from consideration on the final instrument as consensus had been reached (Appendix I). To avoid the mail delays reported earlier by several panelists, nine of the thirteen instruments were sent express mail to allow panelists greater response time.

Again, follow-up postcards were sent one week after the initial mailing thanking those who had returned their instruments and urging those who had not to do so. Follow-up phone calls were made for those instruments not returned by the due date, Wednesday, March 20. Because one instrument was lost, round three was not completed until April 17.

**Data Analysis**

For each item in each round panelists' comments were compiled, the percent of agreement and disagreement and mean certainty score were calculated according to methods prescribed by Warren, Klonglan, and Sabri (1969). By making respondents mark agree or disagree plus the degree of certainty for the item, Warren et al. believed respondents would be forced to think about their response more carefully. They provided a scale (Figure 2) to transform those responses into a more workable form for data analysis.
transformation scale used was chosen because it did not assume equal
intervals between points and weighted the strongest opinions more
heavily.

Disagree 5 4 3 2 1 D/A 1 2 3 4 5 Agree
Transformed Scores 0 3 5 6 7 8 9 10 11 13 16

Figure 2. Transformation of Certainty Scores

Finally, the notes from the researchers in rounds two and three
represented the analysis of the qualitative data and reacted to,
refined or raised further questions about the panelists' comments. A
formative review prior to each round was used to verify or correct for
researcher bias in the analysis and interpretations of the panelists'
comments.
CHAPTER IV

Findings

This study used a modified three-round Delphi technique to answer the following research questions:

1. What do courseware evaluation experts, subject matter experts, instructional designers, and teachers believe are the instructional characteristics critical for promoting achievement in courseware designed for grades 9 - 12?

2. To what extent do courseware evaluation experts, subject matter experts, instructional designers, and teachers agree in their beliefs as to which instructional characteristics are critical for promoting achievement in courseware designed for grades 9 - 12?

3. To what extent do courseware evaluation experts, subject matter experts, instructional designers, and teachers disagree in their beliefs as to which instructional characteristics are critical for promoting achievement in courseware designed for grades 9 - 12?

Procedures

A panel of thirteen experts representing the stakeholders in the courseware evaluation process served on the panel. All panelists were nominated and chosen by an unbiased third party to avoid researcher bias. Panelists described themselves in round one as: three
courseware evaluators, one subject matter expert, seven instructional designers, and two classroom teachers. This distribution was particularly interesting as great care had been taken to ensure equal representation of stakeholders. The thirteen expert panel members are listed in Appendix D.

Panelists were sent three rounds of instruments, designed according to Dillman's total design method (Dillman, 1978) and revised through a formative evaluation prior to each mailing.

In round one, panelists were asked to react to those instructional characteristics (1) that had been suggested by a review of the cognitive psychology literature as likely to promote achievement and (2) that could be manipulated by the instructional designer. Student conditions and learning outcomes (Reigeluth and Merrill, 1979) were omitted from the study. The original list of instructional characteristics likely to promote achievement suggested by the research literature included the seventeen items described in chapter three and listed in the round one instrument (Appendix G).

For each instructional characteristic, panelists were asked (1) to agree or disagree with the critical need for an instructional characteristic, (2) to indicate their degree of certainty, (3) to support their position if they so desired and (4) to suggest additional characteristics. Minimal demographic information was also sought.

Thirteen instruments had been returned by the round one cut-off date, approximately two weeks after the due date. In round one,
panelists suggested five additional instructional characteristics which were added to the item pool for instruments two and three.

By the end of round one, the panel had reached 100% agreement on six items critical for promoting achievement in courseware designed for grades 9-12:

1. opportunities for success,
2. corrective feedback (i.e., explanation of incorrect response),
3. opportunities for practice,
4. opportunities for problem solving,
5. opportunities to elaborate (i.e., further describe or build upon ideas within the lesson), and
6. graphic and visual cues

In rounds two and three, panelists were (1) given the percentage of agreement and disagreement for each item in the previous round; (2) shown their positions from the previous round; (3) shown the comments from the previous round; (4) given notes from the researcher in cases where clarifications, rewordings, or summations were needed; and (5) asked to reconsider their earlier positions.

In round two, each person reconsidered his/her opinion; thus, the instructional characteristic graphic and visual cues, previously eliminated because of earlier consensus, was returned to round three for reconsideration. By the end of round two, the panel had reached 100% agreement on three additional instructional characteristics:

1. learner control of routine study options (e.g., pace of movement through the lesson, help screens, amount of practice, opportunity to review, quit option),
2. opportunities within the program to ask questions of the program, and
3. immediate feedback.

The panel did not reach consensus on additional items in round three.
At the conclusion of the study, panel members were given the opportunity to comment about their experiences as panelists and on implications for further research (Appendix J). Many of the panelists stated they had enjoyed participating in the study and looked forward to receiving the results. Surprisingly, no one complained about the length of the study or potential inconvenience of the Delphi process, although one did mention experiencing frustration trying to interpret the items. Several panelists mentioned a desire to meet fellow panel members or to explore ideas further. Interestingly, two panelists questioned the expert nature of the panel.

Panelists felt further research should (1) further refine some of the instructional characteristics, (2) explore the relationship between the instructional characteristics and specific learning outcomes (e.g., drill, practice, tutorial, or simulation), (3) identify courseware which embodied the eight instructional characteristics on which consensus was reached and (4) explore ways to integrate the courseware into the 9-12 curriculum. As one panelist said, however, "The real challenge is to convince software creators to follow our suggestions."

For each item in each round, the researcher calculated the percentage of agreement and disagreement, determined the mean certainty score (Figure 2), and compiled the comments from the panel. Percentage of agreement and disagreement scores were rounded off to the nearest hundredth and mean certainty scores to the nearest thousandth. Any discrepancies between the percentages reported in the
instruments (Appendixes G, H, I) and Tables 1 and 2 within Chapters IV and V are caused by rounding procedures. In a few instances, a panelist provided a narrative response to the question but felt unable to indicate the degree of certainty. In such cases, the mean certainty score for that item was computed using only the number of scores marked. Such instances are noted in the tables.

<table>
<thead>
<tr>
<th></th>
<th>most</th>
<th>least</th>
<th>least</th>
<th>most</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISAGREE</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>A/D</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>AGREE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Transformed Scores

0 3 5 6 7 8 9 10 11 13 16

Figure 2. Transformation of Certainty Scores

Results

The first research question regarding what courseware evaluation experts, subject matter experts, instructional designers, and teachers believed were the instructional characteristics critical for promoting achievement was answered by achieving consensus on the following instructional characteristics:

1. Learner control of routine study options (e.g., pace of movement through the lesson, help screens, amount of practice, opportunity to review, quit option),
2. Opportunities for success,
3. Corrective feedback (i.e., explanation of incorrect response),
4. Opportunities for practice,
5. Opportunities for problem solving,
6. Opportunities to elaborate (i.e., further describe or build upon ideas within the lesson),
7. Opportunities within the program to ask questions of the program, and
8. Immediate feedback.

All the instructional characteristics have been ranked from the highest to the lowest certainty strength in Table 1. A closer look at how the non-consensus items changed from round to round follows.
<table>
<thead>
<tr>
<th>Instructional Characteristic</th>
<th>n</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Learner control of routine study options (e.g., pace of movement through the lesson, help screens, amount of practice, opportunity to review, quit option)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>13</td>
<td>100</td>
<td>15.153</td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td>0</td>
<td>0</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>15.153</td>
<td></td>
</tr>
<tr>
<td>Opportunities for success</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>13</td>
<td>100</td>
<td>15.076</td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td>0</td>
<td>0</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>15.076</td>
<td></td>
</tr>
<tr>
<td>Corrective feedback (i.e., explanation of incorrect responses)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>13</td>
<td>100</td>
<td>14.538</td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td>0</td>
<td>0</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>14.538</td>
<td></td>
</tr>
<tr>
<td>Opportunities for practice</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>13</td>
<td>100</td>
<td>14.384</td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td>0</td>
<td>0</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>14.384</td>
<td></td>
</tr>
<tr>
<td>Opportunities for problem solving</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>13</td>
<td>100</td>
<td>14.307</td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td>0</td>
<td>0</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>14.307</td>
<td></td>
</tr>
<tr>
<td>Opportunities to elaborate (i.e., further describe or build upon ideas within the lesson)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>13</td>
<td>100</td>
<td>13.384</td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td>0</td>
<td>0</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>13.384</td>
<td></td>
</tr>
<tr>
<td>Instructional Characteristic</td>
<td>n</td>
<td>%</td>
<td>X Certainty</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------------------------------------------------------</td>
<td>----</td>
<td>----</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td><strong>Opportunities within the program to ask questions of the program</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>13</td>
<td>100</td>
<td>13.769</td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td>0</td>
<td>0</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>13.769</td>
<td></td>
</tr>
<tr>
<td><strong>Immediate feedback</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>13</td>
<td>100</td>
<td>13.615</td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td>0</td>
<td>0</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>13.615</td>
<td></td>
</tr>
<tr>
<td><strong>Questions requiring recall, comprehension, analysis, synthesis and/or evaluation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>12</td>
<td>92</td>
<td>13.833</td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td>1</td>
<td>8</td>
<td>3.000</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>13.000</td>
<td></td>
</tr>
<tr>
<td><strong>Opportunities to imitate models</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>12</td>
<td>92</td>
<td>13.166</td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td>1</td>
<td>8</td>
<td>3.000</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>12.384</td>
<td></td>
</tr>
<tr>
<td><strong>On-line option for overview of concepts to be learned</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>11</td>
<td>85</td>
<td>13.818</td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td>2</td>
<td>15</td>
<td>* 0.000</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>*12.666</td>
<td></td>
</tr>
<tr>
<td><strong>Lessons organized for each learner from the least to the most difficult learning task by some mechanism</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>11</td>
<td>85</td>
<td>13.636</td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td>2</td>
<td>15</td>
<td>4.000</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>12.154</td>
<td></td>
</tr>
<tr>
<td><strong>On-line option for overview of skills</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>11</td>
<td>85</td>
<td>13.000</td>
<td></td>
</tr>
<tr>
<td>Disagree</td>
<td>2</td>
<td>15</td>
<td>* 5.000</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>*11.916</td>
<td></td>
</tr>
</tbody>
</table>
Table 1 (continued)

<table>
<thead>
<tr>
<th>Instructional Characteristic</th>
<th>n</th>
<th>%</th>
<th>X Certainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphic and visual cues (e.g., color coding, diagrams, charts)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>11</td>
<td>85</td>
<td>13.272</td>
</tr>
<tr>
<td>Disagree</td>
<td>2</td>
<td>15</td>
<td>4.000</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>11.846</td>
</tr>
<tr>
<td>On-line statement of goals and objectives</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>11</td>
<td>85</td>
<td>12.909</td>
</tr>
<tr>
<td>Disagree</td>
<td>2</td>
<td>15</td>
<td>3.000</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>11.384</td>
</tr>
<tr>
<td>Opportunities for teacher to &quot;author&quot; some input</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>9</td>
<td>69</td>
<td>12.333</td>
</tr>
<tr>
<td>Disagree</td>
<td>4</td>
<td>31</td>
<td>2.750</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>8.769</td>
</tr>
<tr>
<td>Organizers embedded within the text</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>8</td>
<td>62</td>
<td>13.500</td>
</tr>
<tr>
<td>Disagree</td>
<td>5</td>
<td>32</td>
<td>3.600</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>9.692</td>
</tr>
<tr>
<td>Portrays main point visually</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>8</td>
<td>61</td>
<td>12.750</td>
</tr>
<tr>
<td>Disagree</td>
<td>4</td>
<td>31</td>
<td>3.500</td>
</tr>
<tr>
<td>Confused</td>
<td>1</td>
<td>8</td>
<td>* 0.000</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>* 9.666</td>
</tr>
<tr>
<td>Verbal imagery (e.g., metaphors, similes, analogies, descriptive phrases)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>7</td>
<td>54</td>
<td>13.000</td>
</tr>
<tr>
<td>Disagree</td>
<td>6</td>
<td>46</td>
<td>3.833</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>8.769</td>
</tr>
<tr>
<td>Learner control of instructional content</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>7</td>
<td>54</td>
<td>13.428</td>
</tr>
<tr>
<td>Disagree</td>
<td>6</td>
<td>46</td>
<td>2.666</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>8.461</td>
</tr>
</tbody>
</table>
Table 1 (concluded)

<table>
<thead>
<tr>
<th>Instructional Characteristic</th>
<th>n</th>
<th>%</th>
<th>$\overline{X}$ Certainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presentation of real-world circumstances, not contrived information</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>6</td>
<td>46</td>
<td>12.666</td>
</tr>
<tr>
<td>Disagree</td>
<td>7</td>
<td>54</td>
<td>* 3.500</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>* 7.461</td>
</tr>
<tr>
<td>Open-endedness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>6</td>
<td>46</td>
<td>12.000</td>
</tr>
<tr>
<td>Disagree</td>
<td>6</td>
<td>46</td>
<td>2.666</td>
</tr>
<tr>
<td>Confused</td>
<td>1</td>
<td>8</td>
<td>* 0.000</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>* 7.333</td>
</tr>
<tr>
<td>Use of humor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agree</td>
<td>1</td>
<td>8</td>
<td>13.000</td>
</tr>
<tr>
<td>Disagree</td>
<td>12</td>
<td>92</td>
<td>2.545</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>100</td>
<td>3.416</td>
</tr>
</tbody>
</table>

Refinement of Instructional Characteristics

Questions requiring recall, analysis, synthesis, and/or evaluation. Originally worded depth or level of questioning, the characteristic was reworded as stated above and clarified to include both implicit (as when the cursor appears without a stated question) as well as explicit questions.

Several comments rejecting the emphasis on text surfaced throughout the three rounds. For example, "Much excellent software does not question at all; it leads the student to ask questions and provides ways to discover answers to one's own questions." Generally, however, panelists supported the need for some mechanism to elicit
higher levels of cognitive processing: analysis, synthesis, application, or evaluation. "If you want analysis as an outcome, call for it in the process of learning," said one panelist.

Most (92%) agreed with the item as it was finally stated; one (8%) disagreed. No reason was stated for the disagreement in round three. The final mean group certainty score, 13.0, reflected the strong degree of certainty on this item.

Opportunities to imitate models. Opportunities to imitate models was an item on which the panelists did not change their opinion from round to round. Two panelists suggested the importance of model imitation depended on the type of skill being taught; another panelist suggested the importance of model imitation depended on the stage of the learning process. A fourth panelist suggested, "exploration of assumptions behind models is a valuable method of learning complex relationships or elements of systems," but other panelists warned against allowing students to believe that the models were perfect and the only models acceptable.

Throughout the study, 92% of the panel agreed with the item and 8% per cent disagreed. The final mean certainty score was 12.384.

On-line option for overview of concepts to be learned. Originally, this item read "overview of skills or concepts to be learned." At the suggestion of panel the item was broken into separate categories of skills and concepts for round two.

Some panelists suggested that such an overview was better suited (1) for adult learners, (2) with textual materials, or (3) for lesson
designers. One panelist suggested that although data show overviews of concepts may lead to more efficient learning, the need for such an overview may not be critical. Other panelists suggested ways for stating the overview (briefly, and in concrete terms); others reacted to earlier comments that the overview be in the documentation and suggested that this feature must be in the hands of the learner. "We can not assume a teacher will always be making decisions about software. These (concepts) should be included so any learner can use them or decide to bypass them." Several other panelists agreed that an overview of concepts must be an option for students to access. As a result, the item was revised to read an "on-line option for overview of concepts to be learned."

On this item, 85% of the panel agreed that it was critical for promoting achievement and 15% disagreed. The final mean certainty score, 12.666 of a possible 16.000, reflected a strong certainty for this characteristic.

Lessons organized for each learner from the least difficult to the most difficult learning task by some mechanism (e.g., a pre-assessment or internal self-adjustment) within the lesson. This item originally read "increased progression of skills or concepts to be learned." Numerous comments from panelists suggested that although the lesson might be designed from the least to most difficult learning task, students should have the option to approach the material as they desired. Lesson organization "must allow access to learners at
various levels on many different tracks, not just one specific progression." In other words, according to another panelist, "courseware should allow for pre and posttesting and branching so instruction is individualized." Yet, another panelist believed that lesson organization depended upon the content, type of learner, and level of learning outcome desired. This instructional characteristic was considered critical by 85% of the panel and not critical by 15%. The final mean certainty score was 12.154.

**On-line overview of skills to be learned.** As mentioned previously, this item was originally stated as an overview of skills and concepts to be learned. Comments from the panel led the researcher to separate skills and concepts and to force panelists to consider whether or not this should be an on-line option. Many of the same comments made item regarding the importance of an on-line overview of concepts were repeated for this item as well. Most of the panel (85%) agreed with the critical need for an on-line option to overview skills and two (15%) disagreed. The final mean certainty score, 11.916, was lower than the 12.666 certainty score for the need for an on-line overview of concepts.

**Graphic and visual cues (e.g., color coding, diagrams, charts).** The panel reached consensus on this item in round one. Such comments as "most important in initial learning," "a picture can be worth one thousand words," and "the literature supports the use of graphic cues" reinforced the consensus. On the successive iteration, however, one
panelist decided the characteristic was important but not critical for all types of learners or skills, so the item was returned to the item pool for further consideration. By the end of round three, another panelist disagreed with the critical need for graphic and visual cues. "Such cues are useful. Color is attractive and preferred by most students and instructors, but the data never show significant advantages for heavy graphic cues." Thus, by the end of the study, 85% of the panel agreed with the item and 15% disagreed. The final mean certainty score was 11.846.

On-line statement of goals and objectives. The majority of the panelists (85%) considered an on-line statement of goals and objectives critical for promoting achievement, although the final mean certainty score was a moderate 11.384. The original item did not specify that the statement of goals and objectives be on-line, but several panelists believed the use of these goals and objectives should be optional.

One who supported the critical need for an on-line statement of goals and objectives argued that one "can't measure effectiveness unless goals and objectives are stated." Others, however, argued that these goals "may occur after the fact with some excellent materials" or although "the learner must know what to do, this does not always require a strict behavioral statement." Another stated, "inferred goals are often at least as effective as the designer's stated or intended goals."
This item particularly provoked debate about the purpose of courseware. Some panelists felt this characteristic more critical with drill, practice, and tutorials than with simulations. Others thought the best use of the computer was for discovery learning; thus, rendering this instructional characteristic useless. As one panelist commented:

"The best software packages I've seen simply provide fascinating microworlds in which to adventure, explore and learn by trial and error; they motivate by stimulating the student's curiosity. Statements of goals and objectives may be perceived by many students as subtle threats of tests to come. While they (on-line statements) may add to learning, they can also take away from the joy of learning. Good software can, in some cases, make them unnecessary by using the joy of learning as a strong intrinsic motivator."

Opportunities for the teacher to author some input. This panel-generated item recieived 69% agreement and 31% disagreement. The final mean certainty score was 8.769.

Several panelists believed this characteristic would facilitate curricular integration ("the best instruction takes place when an environment matches specific needs of individuals") but others questioned the impact on achievement ("greatly aids local acceptance but not directly necessary for student achievement if teacher is comfortable with the material and sufficiently familiar with it to provide appropriate integration with other instruction.") Dissenters argued, however, "(this) takes control of design and replicability of results out of the designer's hands" and, "this is no more critical with computer software than it is with textbooks."
One panelists stated, "We (teachers) don't have time to do it (author courseware) well. (A better idea would be to) provide opportunities for students to modify and extend the program."

Nevertheless, one must wonder why students would be able to author courseware well if their teachers can not. As one panelist suggested, "student recoding of the program presents some real hazards."

Organizers embedded within the text (e.g., questions, headings, sub-headings). This item focused upon the need for textual organizers. Although panelists believed some type of organizer, either textual or graphic, might be helpful; they generally did not believe textual organizers were critical for promoting achievement. Two panelists argued that good courseware should avoid the use of text. One noted, "Text? If (the) program is mostly text, it should probably be a textbook. The computer is a visual medium." Several panelists believed the importance of this instructional characteristic depended on the type of courseware.

The item received 62% agreement and 38% disagreement with a low, mean certainty score of 9.692.

Portrays main point visually. Proponents of the fundamentally visual nature of this learning technology again strongly supported this position. "Learners will remember the information if it has been portrayed visually. The visuals must be designed to correctly convey the information, not thrown in for decoration." Or, as another
panelist commented, "If the main point is not presented visually and interactively, it (the content) can probably be presented more efficiently by a textbook." On the other hand, those who disagreed tended to think the characteristic useful but not critical ("important in some situations but not critical in all").

At the conclusion of round three 61% agreed, 31% disagreed, and 8% were confused. The final mean certainty score was 9.666.

Verbal imagery (e.g., metaphors, similes, analogies, descriptive phrases). Opinion of this item never changed throughout the study. In every round, 54% of the panel agreed with the item and 46% disagreed. The final mean certainty score of 8.769 supported the wide range of opinion.

One panelist noted that verbal imagery was particularly effective when introducing new concepts, but another believed this type of aid was too dependent upon student experience. One mentioned that "research supports analogies if the object or concept used as the analogy is familiar to the learner." Again, a few panelists objected to this instructional characteristic because it emphasized the textual rather than the more desirable graphic nature of courseware. "Some excellent and effective pieces of software are entirely non-verbal," stated one panelist.

Learner control of instructional content. In round one, this item was part of a more general category, high degree of learner control. Successive iterations refined the category, separating learner control for routine study options (on which consensus was
achieved in round two) and learner control of instructional content. In round three, panelists were told to assume that learner control of instructional content could allow students to bypass parts of the content whether or not competent.

In reaction to this clarification, one panelist suggested such an option might create a problem for students who were not yet independent learners and thus result in inefficient learning. Another panelist believed this type of control "still very much depends on the learner, content, or grade level." As one panelist stated,

"Many proponents of learner control think only of this goal (teaching 9 - 12 students to be independent learners), however, and fail to understand the distinction between preparation for independent learning (mostly non-school or post-secondary) and school-supported learning (which is obligated to provide efficient learning of the basics at low cost.)"

On this item 54% agreed and 46% disagreed. The final mean certainty score, 8.461 of 16.000, reflected wide disagreement on the critical nature of learner control of the instructional content.

Presentation of real-world, not contrived information. Although submitted by a panelist, many participants philosophically questioned the ideas of real and contrived. One panelist even argued, "This statement, if applied to printed materials, would deny the validity and value of all fiction." Another believed the importance of real versus contrived information depended on the instructional strategy and lesson content. The 46% agreement and 54% disagreement as well as the final mean certainty score of 7.461 reflected the split nature of the panels' position on this characteristic.
Open-endedness. Open-endedness, another panel-generated item, never was successfully clarified. Some panelists believed this characteristic was beyond the current state of the technology; another believed this was better done by other media. Those who supported the need for open-endedness said "too much of current cour6ware is correct answer oriented," and "[open-ended responses] allow for higher levels of thinking, something our high school students tend to lack." Those who disagreed noted, "[the] computer is not appropriate for free-form response" and "[I] still think this can be better done by the human touch."

By the conclusion of round three, 46% agreed, 46% disagreed, and 8% were still confused. The final mean certainty score was 7.333.

Use of humor. Use of humor received the strongest disagreement (92%) of any item in the study. Only 8%, or one panelist agreed that it was critical. The final mean certainty score was 3.416. However, one panelist reported research has indicated that humor can increase retention for some subjects. While some other panelists thought "a light approach" or "whimsical approach" might be desirable, most thought this instructional characteristic too dependent on instructional context and learner preference. One panelist commented that "educational software must depend to some extent on repetition and that very few things are funny the second time around."

Field-tested materials which have demonstrated success with the target population. Submitted in round one by a member of the panel, panelists eventually decided that field-tested materials were the
result of a developmental procedure rather than an instructional characteristic as defined by this study. Thus, the item was withdrawn from the item pool at the end of round three.

Although not the domain of this study, numerous panelists commented upon the crucial need for field-tested materials to help ensure quality courseware. For example, "This (field testing) is the single most important feature to look for when purchasing courseware."

Agreement and Disagreement Among Groups

The second and third research questions sought to understand the extent of agreement and disagreement among the four groups of panelists. Mean certainty scores were used to represent the strength of agreement or disagreement.

When the 18 experts were originally selected, the four nomination categories were equally distributed. When the 13 experts who served on the panel verified their areas of expertise in round one, however, the category distribution shifted dramatically. Because the unequal distribution of experts within the groups suggested possible frame error, the degree of agreement and disagreement for each item only has been described. The possible frame error precluded further analysis of these two research questions. Table 2 describes the degree of agreement among groups.
Table 2

**Description of Agreement and Disagreement Among Groups**

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>% Agree</th>
<th>% Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learner control of routine study options</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE</td>
<td>3</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>SME</td>
<td>1</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>ID</td>
<td>7</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>T</td>
<td>2</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Opportunities for success</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE</td>
<td>3</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>SME</td>
<td>1</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>ID</td>
<td>7</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>T</td>
<td>2</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Corrective feedback</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE</td>
<td>3</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>SME</td>
<td>1</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>ID</td>
<td>7</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>T</td>
<td>2</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Opportunities for practice</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE</td>
<td>3</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>SME</td>
<td>1</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>ID</td>
<td>7</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>T</td>
<td>2</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Opportunities for problem solving</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE</td>
<td>3</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>SME</td>
<td>1</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>ID</td>
<td>7</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>T</td>
<td>2</td>
<td>100</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: CE = courseware evaluators, SME = subject matter experts, ID = instructional designers, T = teachers.
Table 2 (continued)

<table>
<thead>
<tr>
<th>Opportunities to elaborate</th>
<th>n</th>
<th>% Agree</th>
<th>% Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>CE 3</td>
<td>100</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>SME 1</td>
<td>100</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>ID 7</td>
<td>100</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>T 2</td>
<td>100</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Opportunities within the program to ask questions of the program</th>
<th>n</th>
<th>% Agree</th>
<th>% Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>CE 3</td>
<td>100</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>SME 1</td>
<td>100</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>ID 7</td>
<td>100</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>T 2</td>
<td>100</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Immediate feedback</th>
<th>n</th>
<th>% Agree</th>
<th>% Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>CE 3</td>
<td>100</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>SME 1</td>
<td>100</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>ID 7</td>
<td>85</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>T 2</td>
<td>100</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Questions requiring recall, analysis, synthesis, and/or evaluation</th>
<th>n</th>
<th>% Agree</th>
<th>% Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>CE 3</td>
<td>100</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>SME 1</td>
<td>100</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>ID 7</td>
<td>85</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>T 2</td>
<td>100</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Opportunities to imitate models</th>
<th>n</th>
<th>% Agree</th>
<th>% Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>CE 3</td>
<td>100</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>SME 1</td>
<td>100</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>ID 7</td>
<td>85</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>T 2</td>
<td>100</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

Note: CE = courseware evaluators, SME = subject matter experts, ID = instructional designers, T = teachers.
Table 2 (continued)

<table>
<thead>
<tr>
<th>Feature Description</th>
<th>CE</th>
<th>SME</th>
<th>ID</th>
<th>T</th>
<th>% Agree</th>
<th>% Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>On-line option for overview of concepts to be learned</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>66.6</td>
<td>33.3</td>
</tr>
<tr>
<td>Lessons organized for each learner from the least difficult to the most difficult learning task by some mechanism (e.g., a pre-assessment or internal self-adjustment) within the lesson</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>On-line overview of skills to be learned</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>66.6</td>
<td>33.3</td>
</tr>
<tr>
<td>Graphic and visual cues (e.g., color coding, diagrams, charts)</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>On-line statement of goals and objectives</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Opportunities for the teacher to author some input</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>66.6</td>
<td>33.3</td>
</tr>
</tbody>
</table>

Note: CE = courseware evaluators, SME = subject matter experts, ID = instructional designers, T = teachers.
Table 2 (continued)

<table>
<thead>
<tr>
<th>Organizers embedded within the text</th>
<th>CE</th>
<th>3</th>
<th>66.6</th>
<th>33.3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SME</td>
<td>1</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>ID</td>
<td>7</td>
<td>71</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>2</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Portrays main point visually</th>
<th>CE</th>
<th>3</th>
<th>66.6</th>
<th>33.3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SME</td>
<td>1</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>ID</td>
<td>6</td>
<td>66.6</td>
<td>33.3</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>2</td>
<td>100</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Verbal imagery (e.g., metaphors, similes, analogies, descriptive phrases)</th>
<th>CE</th>
<th>3</th>
<th>33.3</th>
<th>66.6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SME</td>
<td>1</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>ID</td>
<td>7</td>
<td>57</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>2</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Learner control of instructional content</th>
<th>CE</th>
<th>3</th>
<th>33.3</th>
<th>66.6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SME</td>
<td>1</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>ID</td>
<td>7</td>
<td>71</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>2</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

| Presentation of real-world, not contrived information | CE | 3 | 100 | 0   |
|-------------------------------------------------------|----|----|-----|
|                                                        | SME| 1  | 100 | 0   |
|                                                        | ID | 6  | 33.3| 66.6|
|                                                        | T  | 2  | 0   | 100 |

Note: CE = courseware evaluators, SME = subject matter experts, ID = instructional designers, T = teachers.
Table 2 (concluded)

<table>
<thead>
<tr>
<th>Open-endedness</th>
<th>CE</th>
<th>66.6</th>
<th>33.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>SME</td>
<td>1</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>ID</td>
<td>6</td>
<td>33.3</td>
<td>66.6</td>
</tr>
<tr>
<td>T</td>
<td>2</td>
<td>100</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Use of humor</th>
<th>CE</th>
<th>33.3</th>
<th>66.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>SME</td>
<td>1</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>ID</td>
<td>7</td>
<td>15</td>
<td>85</td>
</tr>
<tr>
<td>T</td>
<td>2</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

Note: CE = courseware evaluators, SME = subject matter experts, ID = instructional designers, T = teachers.
CHAPTER V
Summary, Implications, and Recommendations

In summary, this study used a modified three-round Delphi technique to answer the following research questions:

1. What do courseware evaluation experts, subject matter experts, instructional designers, and teachers believe are the instructional characteristics critical for promoting achievement in courseware designed for grades 9 - 12?

2. To what extent do courseware evaluation experts, subject matter experts, instructional designers, and teachers agree in their beliefs as to which instructional characteristics are critical for promoting achievement in courseware designed for grades 9 - 12?

3. To what extent do courseware evaluation experts, subject matter experts, instructional designers, and teachers disagree in their beliefs as to which instructional characteristics are critical for promoting achievement in courseware designed for grades 9 - 12?

Procedures
Nominations of experts to serve on the Delphi panel were sought from the leaders of four groups considered stakeholders in the courseware evaluation process: courseware evaluators, subject matter experts, instructional designers, and classroom teachers. Twenty-eight persons were nominated as experts. From that group, 18
were systematically selected to serve on the expert panel. A third party was used to select the panelists to avoid researcher bias. Of the original 18 panel members chosen, 13 participated. All of those 13 panelists participated in all three rounds of the study.

The Delphi instruments were designed according to Dillman's total design method (Dillman, 1978) and were revised through formative evaluation prior to each mailing. In round one, panelists were asked to react to those instructional characteristics (1) that had been suggested by a review of the cognitive psychology literature as likely to promote achievement and (2) that could be manipulated by the instructional designer. For each instructional characteristic, panelists were asked (1) to agree or disagree with alternative positions, (2) to indicate their degree of certainty, (3) to support their position if they so desired and (4) to suggest additional characteristics. Minimal demographic information was also sought. Follow-up phone calls were made to secure missing questionnaires. Data analysis included (1) calculation of agreement and disagreement percentages, (2) calculation of mean certainty scores, and (3) compilation of panelists comments.

In rounds two and three, panelists were (1) given the percentage of agreement and disagreement for each item in the previous round; (2) shown their positions from the previous round; (3) shown the comments from the previous round; (4) given notes from the researcher in cases where clarifications, rewordings, or summations were needed; and (5) asked to reconsider their earlier positions. Follow-up postcards were
sent one week after each mailing in rounds two and three, and follow-up phone calls were made for those instruments not returned by the due dates.

In round three, panelists were given the opportunity to comment about their experience as a panelist and on implications for further research. Those comments and the data were tabulated and summarized. (Appendix J)

The first research question regarding what courseware evaluation experts, subject matter experts, instructional designers, and teachers believed were the instructional characteristics critical for promoting achievement was answered by achieving consensus on the following items:

1. learner control of routine study options (e.g., pace or movement through the lesson, help screens, amount of practice, opportunity to review, quit options),
2. opportunities for success,
3. corrective feedback (i.e., explanation of incorrect responses),
4. opportunities for practice,
5. opportunities for problem solving,
6. opportunities to elaborate (i.e., further describe or build upon ideas within the lesson), and
7. opportunities within the program to ask questions of the program.
8. immediate feedback.

The findings on the second and third questions regarding the extent of agreement and disagreement among groups were far less clear. The shifts in the nomination categories during the panel verification process, which suggested possible frame error, indicated that a discussion of the conclusions and implications should be limited to the extent of panel agreement or disagreement. This should not limit
the usefulness of the study because of the expert status and diverse nature of the Delphi panel.

The quality of discussion generated and the outcomes of the study supported the utility of the Delphi technique for bringing consensus to a complex field marked by rapid change. As expected from other Delphi studies (Hostrop, 1975; Sutphin, 1981) the greatest movement in panel opinion occurred between rounds one and two; therefore, even though the items added by panelists in round one underwent only two, rather than three iterations, panelists' positions probably would not have changed as the result of another round. To further refine the issues one might consider using the Delphi or, as one panelist suggested, other techniques such as critical incidence. Nevertheless, the results of this Delphi study suggested some important implications for courseware design, evaluation, curricular integration, and educational organization.

Discussion and Implications of Findings

Findings of this study and their implications will be discussed in several major areas: the panel's confirmation of the findings suggested in the research literature, the need to make explicit and understand the definitions of the evaluation criteria, the importance of specifying the target audience, the consequences of the debates regarding the purpose of courseware and the nature of the medium, and the impacts of individualized instruction.
Confirmation of the Research Literature

First, this study has confirmed (1) the findings reported in the psychology literature and (2) the efficacy of using a modified Delphi based upon such a review. Of the 17 instructional characteristics used as the basis of this study, all but two (verbal imagery and use of humor) elicited 85% agreement or higher. Of the five additional instructional characteristics suggested by panel members in round one, four of those were ranked in the bottom third of the rankings and one was eliminated because it did not meet the definition of instructional characteristic as defined by this study. This suggested that the initial literature review had reported the major characteristics likely to promote achievement. Thus, though the modified Delphi allowed panelists to compensate for oversights in the literature review, there were in fact no significant instructional characteristics added by the panel.

Several possibilities may account for the panel's confirmation of the research literature. One strong possibility is that although many have contended that computer technology offers new and fundamentally different techniques for instruction, the strong agreement on 15 of the 17 instructional characteristics supported the belief that quality instruction should be based on an understanding of how people learn rather than on the characteristics of the instructional medium. As a result, panelists confirmed those instructional characteristics likely to promote achievement rather than emphasizing the technical characteristics of the computer; consequently, findings from research
in cognitive and developmental psychology became the basis for the panel's reaction as panelists cited research findings to support their arguments throughout the study. Even though it is difficult to ascertain why people believe as they do in mailed Delphi rounds, it seems reasonable to conclude that such support carried much weight in the panel's decision-making process.

Another reason for the confirmation of the research literature may have been that the instructional characteristics were not operationally defined and, therefore, may have been too general to have provoked greater disagreement. For example, items such as opportunities for success or opportunities for problem solving were difficult to disagree with as stated. Had the study attempted to define the types of opportunities which provide successful experiences or the kinds of problem-solving activities, panelists most surely would have disagreed with one another, especially had the study attempted to evaluate the relationship between instructional strategy and instructional characteristics. Several panel members commented upon the limited usefulness of such broad instructional characteristics at the conclusion of the study.

Whatever the reason for the confirmation of the research literature, EPIE and other courseware evaluation services have shown that courseware designed for grades 9 - 12 presently does not embody even the general instructional characteristics agreed to by the Delphi panel. Therefore, the results of this study suggest a clear and powerful agenda for future courseware development. If courseware
developers and publishers want to develop quality products, they need to begin incorporating the findings of the research literature, especially in the areas recommended by the panel. Educators and the taxpayers have already begun to demand that publishers provide evidence that the courseware promotes learning rather than simply provides entertainment.

Having suggested the need for more systematic courseware development rooted in research, however, one must consider that the average educational software package presently sells for $42.05 ("Charting the Course", 1985). Because implementation of the instructional characteristics confirmed in this study would require development teams comprised of individuals with considerable expertise in the content area, instructional design, computer graphics and programming; designing quality courseware which realizes a profit becomes a tremendous problem for the courseware industry. Unless better strategies can be developed to protect software from unlawful copying or better licensing agreements can be structured, courseware publishers are likely to continue catering to the more profitable and far less demanding home market; and, thus, avoid implementing the more difficult instructional characteristics suggested by research. That being the case, government or private research projects would be left to develop the more sophisticated programs. Considering the current dismantling of the National Institute of Education, which has funded several projects aimed at improving courseware quality, the immediate
fate of research-based courseware development does not seem encouraging.

**Importance of Defining Evaluation Criteria**

Second, by struggling to understand and then refine the broad instructional characteristics in this three-round Delphi study, the panelists illustrated how difficult but critical it is to communicate a clear definition of criterion. For example, corrective feedback, a consensus item in this study and a term frequently used by courseware evaluators, was specifically defined for the panelists as the explanation of incorrect responses. In many evaluation checklists or services, however, corrective feedback means simply telling the learner whether or not a response was accurate.

Obviously, unless those evaluating a piece of courseware and those reading the evaluation report are using the same definition for a given criteria, the evaluation of that characteristic is meaningless; for this reason, the more sophisticated level one evaluation efforts train their evaluators to use their checklists. The Alberta project considers this training so critical that the teacher evaluators are paid for their training and professional judgments.

Unfortunately, unless specifically trained to use a given set of criteria, few teachers or schools take the time to ensure that all who evaluate courseware are working from the same definitions. For example, the popular MicroSIFT evaluation guidelines regularly are used without any reference to the extensive definitions and
illustrations of the criteria discussed in the booklet which accompanies the guidelines.

Almost every instructional characteristic in this study required further definition and, while forcing such refinement was an integral part of the Delphi process, most existing evaluation checklists consist of broad criteria. Without a clear understanding of the meaning of those criteria through systematic instrument development, validation, and user training; evaluations based upon checklists and guidelines fail to be useful. Ironically, using courseware evaluation checklists, which have been designed specifically to help simplify a complex process, may actually be confounding rather than clarifying the evaluation process.

For this reason, future evaluation efforts need to move more in the directions suggested by Kansky et al. (1981) and Della-Piana & Della-Piana (1982) which study how the student proceeds through a given lesson and, therefore, make explicit the inherent complexity of courseware evaluation process. Such efforts to study these student paths through the lesson also would emphasize the need for more research aimed at improving our understanding of individual differences.

The proposed new directions in courseware evaluation and research raise some serious ethical questions however. For example, to follow a student's path through a lesson, with or without his/her knowledge, allows one an intimate knowledge of the workings of the learner's mind. Such knowledge, the ultimate invasion of privacy, obviously
places the learner in an untenable position. Renewed discussions regarding the ethical consequences of such recordkeeping and tracking must begin immediately and policy statements must be developed and implemented.

The Need for a Clear Definition of the Target Audience

Third, many of the characteristics endorsed by the Delphi panel relied to some extent upon a clear definition of and a match with the target audience (e.g., opportunities for success, or corrective feedback). Whereas a teacher can readily adapt materials or instructional strategies to accommodate different students, the need to specify the target audience becomes a critical issue for courseware which can not adapt so readily. All possible interactions and adaptations within courseware must be planned during development and verified when field tested. Because most of the 10,000 instructional software programs on the market lack clearly defined target audiences (Bialo & Erickson, 1984), evaluators and teachers have great difficulty identifying courseware which provides appropriately challenging experiences and diagnostic feedback for a specific ability or grade level. Courseware marketed for K-adult, or from crib to grave, may have a wide commercial market but is instructionally inadequate and indefensible. Because courseware evaluation requires knowing the program's designated target audience, educators must demand that courseware publishers identify the target audience more specifically. While courseware evaluation might further indicate
audience suitability, the general nature of the
target audience first needs to be specified by publishers and
developers.

**Consequences of the Debates Regarding the Purpose of Courseware and the Nature of the Medium**

Fourth, two debates raged throughout the study which supported Roblyer's explanations for the lack of quality courseware. The first was the debate about the best use for the computer. Some panelists persisted in resisting any notion of structured teaching through courseware. These panelists strongly favored problem solving and discovery learning courseware. Interestingly, the panel strongly favored problem solving (100% agreement) and model imitation (92%) despite the fact that this study did not focus on just one instructional strategy, such as simulation. Such reaction portends a movement away from the current heavy emphasis on drill and practice and toward greater emphasis on problem solving.

The second debate, somewhat related to the first, concerned the extent to which instructional courseware should rely upon graphic rather than textual presentations. Two panelists particularly resisted the use of text in courseware. Part of their resistance may have been a reaction to early courseware which presented pages of scrolling text, which was often difficult to read and simply replicated textbooks. Another reason for their resistance to text may simply have been the emphasis upon text in the wording of many of the
items in round one. Generally, however, these panelists believed that the computer is primarily a visual medium and as such should rely primarily upon use of graphics to stimulate learning.

While the other panelists strongly agreed that the effective use of graphics was often very effective, graphics were not considered critical for promoting achievement. The popularity of arcade games, which depend on sophisticated graphics, and student preference for programs which include graphics certainly emphasizes the need to better understand the relationship between graphics and cognitive processing. Nevertheless, the graphic and gaming formats for courseware advocated by some panelists raise major problems for the curricular integration of such courseware in schools criticized for their failure to teach basic skills in a traditional manner.

Implications of Individualized Instruction

Fifth, and most importantly, four items on which there was strong agreement focused on the importance of learner control: (1) control of routine study options, 100%; (2) opportunities within the program to ask questions of the program, 100%; (3) the on-line option for overview of concepts to be learned, 85%; and, (4) the on-line option for overview of skills, 85%. Had other items been defined operationally (e.g., opportunities for problem solving), they too might have required learner control.

These findings, which emphasize the need for certain types of learner control, have serious implications for courseware delivery,
curricular structure and integration, teacher training, and the organization of schools. By its very nature, learner control implies individualized instruction. To adapt to the varying needs or desires of the learner, courseware will require sophisticated assessment tools or internal adaptive mechanisms. Consequently, such tools will require sophisticated programs and amounts of memory beyond what most schools presently can or are willing to afford. Unless the schools embrace more sophisticated delivery systems, however, educational programs will continue not to provide the level of learner control endorsed by the Delphi panel.

Should such learner control of instructional courseware ever become a reality, the curricular implications would be profound. The tremendous changes brought on by the influx of paperbacks into the 9-12 curriculum would seem trivial compared to those brought on by the influx of quality courseware. New, more complex curricular and organizational structures would have to be devised to provide for such instructional diversity. Organizational roles would need to be redefined; teachers and administrators would need to be retrained. Schools as they are now organized would become unmanageable.

Other than the critical organizational implications, findings regarding learner control also have particular relevance for the secondary students who would be using the courseware. Because achieving independence is a critical developmental task for adolescents, the panels' emphasis on instructional characteristics supporting learner control most surely supports the relationship
between instructional design and developmental psychology. If courseware were available which allowed greater learner control, secondary school students could assume greater responsibility for their progress and thus become more independent learners. By assuming responsibility for learning, students would probably improve their achievement as well.

The fact that the instructional characteristic allowing the learner, whether or not competent, control of the instructional content received only 54% agreement may reflect the panel's concern about giving teenagers control they were not ready to handle. Interestingly, EPIE's most recent work (although heavily emphasizing K-8 software) reported learner control of even routine study options in only one-half of the programs reviewed.

Because efforts to implement individualized instruction have been so unsuccessful, the existing courseware is so instructionally inadequate, and student access to computers in school has been so limited (30 minutes on the computer per week per student is considered a luxury); it is difficult to imagine the consequences of implementing courseware which embodies the panel's recommendations. Nevertheless, because the number of microcomputers continues to increase and courseware probably will not remain in its primitive state for long, the challenge to traditional classroom and school organization may arise almost by accident within the next five to ten years. This challenge will be confounded by the fact that although only an estimated ten percent of the school districts in the United States
have policies to guide the implementation of computers, there seems to be little effort to develop them. As one superintendent said, "We probably do need policies but the panic is over. Now that buildings have a few computers, parents have settled down and we can get back to business." Ironically, the demand for and the random influx of high quality courseware into the curriculum may pose for administrators the most serious threat to traditional educational organization yet because such courseware finally makes possible individualized instruction throughout the curriculum.

Whatever the impact of courseware, educators will most surely be forced to learn more about the teaching/learning process. Courseware design forces one to analyze instruction far more objectively than ever before; field testing courseware allows one to revise and then replicate an instructional sequence. Such analysis, revision, and replication will force us to extend our understanding of the learner and the learning process.

Implications for Research

Additional research is indicated in several key areas. Educators need to better understand, for example, the relationship between instructional characteristics and instructional strategies. Throughout the study, panelists said that the importance of the instructional characteristic depended upon the instructional strategy employed by the courseware. Although rigid classification of courseware may bring about what Merrill calls "the hardening of the
instructional categories," courseware should be evaluated according to its primary instructional purpose. Consequently, the importance of instructional characteristics may vary depending upon that primary instructional purpose. Research needs to further investigate this relationship.

Further research is needed in order to operationally define the broad instructional characteristics agreed to by the panel. Developers need better data regarding the type of corrective feedback which is most helpful for secondary students. How much practice do these students need? When are graphic and visual cues most helpful?

Courseware developers and evaluators also need to better understand different ways students proceed through a given piece of courseware, despite the serious implications of the related ethical issues. Through a better understanding of the learner's movement through the lesson, developers and evaluators might also better understand the importance of cognitive style and the nature of cognitive processing.

Teacher educators need to discover better ways to train teachers and administrators to evaluate and integrate courseware into the curriculum. Related to this, researchers need to begin systematically exploring the organizational and management implications of computer implementation.
Recommendations

Several recommendations should be made as the result of this study. First, because the expert's conclusions support the research literature, courseware developers should incorporate those findings which finally provide the shared experiences and shared language advocated by Olds (1983) and Stufflebeam (1980). To accomplish this, courseware publishers must rely upon teams of developers which include instructional designers knowledgeable about the research.

Second, courseware publishers need to more specifically define the intended target audience for a given piece of courseware thereby strengthening the development process and enabling educators to further evaluate the courseware and implement it.

Third, schools need to purchase more sophisticated computer delivery systems if they want to use more sophisticated courseware. Sophisticated courseware eventually will require large amounts of memory (especially for programs which rely upon data bases), numerous printer configurations, and computer networks within schools and among districts. Most sophisticated delivery systems will require creative, new ways of financing such systems and new ways of justifying the cost of those expenses. Unfortunately, the latter is particularly difficult to do, even within the business environment. Fourth, the public and private sectors must work together to support more sophisticated and research-based courseware development projects if the computer is to be used as a significant instructional medium. We can not expect publishers to continue charging $42.05 for
sophisticated quality packages — quality courseware will cost much more. Furthermore, schools must buy more than one copy of courseware if that courseware is ever to be integrated into the basic instruction.

Fifth, schools must develop and rely upon policies for computer implementation. Unless educators make explicit the political issues and struggle to solve the difficult implementation questions required in such policy formulation, the influx of computers will result in curricular and organizational chaos.
REFERENCES


APPENDIX A

Correspondence with Nominators of the Expert Panel
Dear:

Numerous evaluation checklists and services have been designed to help teachers identify quality courseware. Unfortunately, we still do not have agreement on the elements most likely to prove instructionally significant. We are proposing, therefore, a national study to help clarify this issue and help your organization better serve the needs of your members or clients.

Because your organization deals with these problems of courseware quality and has a broad knowledge of the qualifications of people in your field, we are asking you to take a few minutes to nominate two people most qualified to serve on a panel of experts to reach agreement on this issue. Your nomination may be made by:

1. considering the selection criteria on the enclosed form,
2. reviewing the enclosed synopsis if you need additional information about the study, and
3. completing and returning the enclosed form by May 4, in the self-addressed stamped envelope.

Your nomination will be one of 20 sought to select this expert panel; consequently, your nominations are critically important. On the enclosed form, we also have requested your permission to notify the experts of your nomination and, thus, encourage their participation in this study.

If you have further questions, please feel free to call us at (614) 422-0711 or 252-0531. We thank you for your assistance and enclose a pen as a small token of our appreciation.

Sincerely yours,

Keith A. Hall

Sallie J. Sherman
December 26, 1984

Dear :  

Last spring you nominated candidates for a national study designed to further investigate the critical instructional factors necessary for quality courseware. From those nominated we have systematically chosen 18 to serve on the expert panel. They will be notified after the first of the year.

We again want to thank you for your nominations and your willingness to assist us.

Sincerely,

Sallie J. Sherman

Keith A. Hall
APPENDIX B

Expert Selection Criteria and Nomination Ballot
SELECTION CRITERIA AND NOMINATION BALLOT

Selection Criteria for "Experts"

An "expert" for the purpose of this study shall be one who:

1. has extensive experience in your field and computer-based education, especially in courseware design and/or evaluation,
2. will be willing to consider alternatives,
3. will be able to construct logically sound deductions of alternate positions which should be accepted with regard to which elements of courseware are most likely to be instructionally significant,
4. has a knowledge of the present state of courseware and will be able to envision the future of courseware development.
5. has a thorough knowledge of the needs and abilities of students, grades 9-12.

Nominations

Please write the names and addresses of your nominees in the space below and return this nomination in the enclosed, self-addressed, stamped envelope by May 4, 1984.

1. Name _______________________________________________
   Address ____________________________________________
   ______________________________________
   ______________________________________

2. Name _______________________________________________
   Address ____________________________________________
   ______________________________________
   ______________________________________

You may ( ) may not ( ) use my name when contacting the nominees.

Signature

Thank you for your assistance.
APPENDIX C

Synopsis Used in Nomination Mailing
SYNOPSIS

INTRODUCTION

The rapid influx of microcomputers into American education has caused many difficult problems for classroom teachers trying to identify quality instructional materials. Often overwhelmed by the technology itself and the exponential growth of the courseware market, educators have turned to professional and commercial guidelines and services for help in identifying quality courseware. Although these guidelines and services have helped, it has been estimated that there now may be as many as 200 evaluation checklists available with others being developed daily. This proliferation of evaluation services, while symptomatic of a complex and growing field, has the potential to confuse rather than enlighten classroom teachers. The purpose of this study, therefore, is to reach agreement regarding instructionally significant elements of courseware. Findings of this sort predictably will bring further clarity to the field and provide a direction for future research.

Methodology and Findings

The study will use a modified Delphi Technique to identify perceptions of a panel of experts nominated by leaders from such fields as: courseware evaluation, academic disciplines, instructional design, and teaching.

The results of this study, targeted on courseware designed for grades 9-12, will be made available to organizations through a final report, presentations at professional meetings, and journal articles.
APPENDIX D

List of Expert Panel Members
List of Expert Panel Members

Dr. Kathryn Alesandrini
Professor, Department of Secondary Education
California State University

Mr. Bruce Brombacher
Teacher
Upper Arlington, OH

Ms. Marlea Jordan
Resource Teacher - CAI/Reading and Math
Dayton, OH

Mr. Daniel West
Teacher
University of Chicago Laboratory Schools

Dr. Gabriel Della-Piana
Professor, Department of Educational Psychology
University of Utah

Ms. Beverly Sangston
Director, Department of Computer-Related Instruction
Montgomery County Public Schools, Maryland

Mr. Donald Holznagel
Director, Computer Technology Program
Northwest Regional Laboratory

Dr. Thomas Liao
Director, Microcomputer Laboratory
SUNY at Stony Brook

Dr. Allen Avner
Director of Evaluation, CERL
University of Illinois

Dr. Peggy Roblyer
Director, ICON Enterprises
Tallahassee, FL

Dr. Robert Caldwell
President
PANDA Learning Systems
APPENDIX E

Correspondence with Expert Panel
Dear:

Numerous evaluation procedures have been designed to identify high-quality courseware. Unfortunately, despite almost 200 evaluation checklists and services, we have little agreement about which characteristics are most likely to promote achievement.

You have been nominated to serve on an expert panel in a national study designed to help arrive at such agreement. Editors of professional computing magazines, directors of software clearing houses, presidents of professional organizations, and leading software publishers nominated you and 27 others to serve on the expert panel. From that original group a systematic method was used to select the present 18 member panel.

The purpose of this study is to reach agreement regarding which instructional characteristics are most likely to promote achievement in courseware designed for secondary school students. Findings of this sort predictably will bring further clarity to the field and provide a direction for future research.

We will be using a modified, three round Delphi technique and your participation is requested in all three rounds. We would like your responses to round one by Wednesday, January 9.

We look forward to working with you on this project and have enclosed a disk holder to thank you for your participation.

Sincerely,

Sallie J. Sherman  
Keith A. Ball
February 17, 1985

Dear:

Excellent comments from you and your colleagues in round one of our Delphi study should make this round thought provoking, although hopefully not too time consuming. We look forward to your reactions and value your comments.

Although we realize that you all are incredibly busy, this study must be finished by the end of March and we still have one more round to complete. Please take some time now to react to the instrument. We need your responses by Wednesday, February 27.

Enclosed is the instrument, a stamped return envelope, and a small token of our appreciation.

Again, we value your participation and thank you for your excellent comments.

Sincerely,

Sallie J. Sherman          Keith A. Hall
March 10, 1985

Dear :

Enclosed is the final instrument of our national Delphi study. Because the panel reached consensus on six items in rounds one and two, you should find this instrument a bit shorter - although not necessarily less thought-provoking. We have deleted, however, all exclamations and underlining from your comments to preserve the "reasoned professional judgment" sought from a Delphi study.

For those of you planning to attend AERA in Chicago, the results of this study will be presented on Monday, April 1 (we hope the date is not significant) at 8:15 a.m. We hope to see some of you then, in some of the other sessions, or at the OSU reception. Please stop and say "hello."

Enclosed, once again, is a small token of our appreciation and a stamped return envelope. We need your responses by Wednesday, March 20. Please try to help us meet our deadline.

Thank you again for your time and insights. You have been a pleasure to work with. We hope that this study has been as challenging for you as it has for us, and that the field will be stimulated by the results.

Sincerely,

Sallie J. Sherman                 Keith A. Hall
APPENDIX F

Pilot Test Instrument Review Panels
Round One Instrument Review Panel
Frances Blake
Sally Hall
Larry Miller
Delia Neuman
Mark Tisone
Eric Watkins

Round Two Instrument Review Panel
Frances Blake
Larry Miller
Delia Neuman
Eric Watkins

Round Three Instrument Review Panel
Frances Blake
Delia Neuman
Robin Riggs
APPENDIX G

Round One Delphi Instrument
Instructional Characteristics Critical For Promoting Achievement In Courseware Designed For Grades 9-12

National Delphi Study

The Ohio State University
College of Education
Instructional Design and Technology
216 Ramseyer Hall
29 West Woodruff Avenue
Columbus, OH 43210
INSTRUMENT #1

INTRODUCTION

This instrument contains two parts. Part one asks you to agree or disagree with alternative positions, to indicate your degree of certainty, and to support your position if you so choose. Part two seeks minimal, essential demographic data.

Only those instructional characteristics (1) that have been suggested by a review of the cognitive psychology literature as likely to promote achievement and (2) that can be manipulated by the instructional designer have been included. Student characteristics (e.g., readiness, cognitive style, intrinsic learner motivation) and implementation concerns have been omitted from this study. The specific questions to be investigated include:

1. What do courseware evaluation experts, subject matter experts, instructional designers, and teachers believe are the instructional characteristics most likely to promote achievement?

2. To what extent do courseware evaluation experts, subject matter experts, instructional designers, and teachers agree in their beliefs as to which instructional characteristics are most likely to promote achievement?

3. To what extent do courseware evaluation experts, subject matter experts, instructional designers, and teachers disagree in their beliefs as to which instructional characteristics are most likely to promote achievement?

We will be using a modified, three-round Delphi Technique. In round one you will be stating your positions; in rounds two and three you will be comparing your positions with the positions of other panel members. Because one of the critical features of this technique is anonymity, you are asked not to discuss this study while it is in progress, especially with other members of the panel. You are, however, welcome to consult whatever data you feel might be helpful. Your answers will be completely confidential. Your questionnaire has been coded so we can verify returns. We would like to receive your responses by Wednesday, January 9.
PART ONE — Alternative Positions

Directions

As you consider each of the characteristics listed, please do the following:

1. Check whether you agree or disagree that the instructional characteristic is critical for promoting achievement in courseware designed for grades 9-12.

2. Using the following descriptions, circle the number which best describes the certainty of your answer.

   Least Certain Somewhat Certain Very Certain Most Certain
   1 2 3 4 5

3. Optional—Use the space provided to support your position.

Example

ISSUE Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

(Check "agree" or "disagree" and circle level of certainty.)

Q-1. attractiveness of packaging

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>DISAGREE</td>
<td>LEAST MOST</td>
</tr>
</tbody>
</table>

(The respondent disagrees with characteristic's instructional importance. It is most certain of their response.)

Packaging is a marketing concern not an instructional design concern.

page 3
**Issue**: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

(Check "agree" or "disagree" and circle level of certainty.)

<table>
<thead>
<tr>
<th>1. statement of goals and objectives</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AGREE</strong></td>
</tr>
<tr>
<td>CERTAINTY</td>
</tr>
<tr>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td><strong>DISAGREE</strong></td>
</tr>
<tr>
<td>LEAST</td>
</tr>
<tr>
<td>MOST</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. increased progression of skills or concepts to be learned</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AGREE</strong></td>
</tr>
<tr>
<td>CERTAINTY</td>
</tr>
<tr>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td><strong>DISAGREE</strong></td>
</tr>
<tr>
<td>LEAST</td>
</tr>
<tr>
<td>MOST</td>
</tr>
</tbody>
</table>
Issue: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

(Check "agree" or "disagree" and circle level of certainty.)

3. use of humor

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DISAGREE</th>
<th>LEAST</th>
<th>MOST</th>
</tr>
</thead>
</table>

4. depth or level of questioning

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DISAGREE</th>
<th>LEAST</th>
<th>MOST</th>
</tr>
</thead>
</table>
Issue: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

(Check "agree" or "disagree" and circle level of certainty.)

5. overview of skills or concepts to be learned

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>DISAGREE</td>
<td>LEAST  MOST</td>
</tr>
</tbody>
</table>

6. opportunities within the program to ask questions of the program

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>DISAGREE</td>
<td>LEAST  MOST</td>
</tr>
</tbody>
</table>
Issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

(Check "agree" or "disagree" and circle level of certainty.)

7. organizers embedded within the text (e.g., questions, headings, sub-headings)

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DISAGREE LEAST MOST

8. graphic or visual cues (e.g., color coding, diagrams, charts)

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DISAGREE LEAST MOST
**Issue** Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

(Check "agree" or "disagree" and circle level of certainty.)

<table>
<thead>
<tr>
<th></th>
<th>AGREE</th>
<th>CERTAINTY</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. verbal imagery (e.g., metaphors, similes, analogies, descriptive phrases)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>AGREE</th>
<th>CERTAINTY</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. opportunities to imitate models</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

page 8
**Issue** Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

(Check "agree" or "disagree" and circle level of certainty.)

<table>
<thead>
<tr>
<th></th>
<th>AGREE</th>
<th>CERTAINTY</th>
<th>DISAGREE</th>
<th>LEAST</th>
<th>MOST</th>
</tr>
</thead>
<tbody>
<tr>
<td>11. knowledge of results</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. corrective feedback</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

(Check "agree" or "disagree" and circle level of certainty.)

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>opportunities for practice</strong></td>
<td><strong>AGREE</strong></td>
<td><strong>CERTAINTY</strong></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td><strong>DISAGREE</strong></td>
<td><strong>LEAST</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>MOST</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>high degree of learner control</strong></td>
<td><strong>AGREE</strong></td>
<td><strong>CERTAINTY</strong></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td><strong>DISAGREE</strong></td>
<td><strong>LEAST</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>MOST</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

(Check "agree" or "disagree" and circle level of certainty.)

15. opportunities for problem solving

- **AGREE**
  - CERTAINTY
    - 1
    - 2
    - 3
    - 4
    - 5

- **DISAGREE**
  - LEAST
  - MOST

16. opportunities to elaborate
   (i.e., further describe or build upon ideas within the lesson)

- **AGREE**
  - CERTAINTY
    - 1
    - 2
    - 3
    - 4
    - 5

- **DISAGREE**
  - LEAST
  - MOST
Issue: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

(Check "agree" or "disagree" and circle level of certainty.)

17. opportunities for success

<table>
<thead>
<tr>
<th>Agree</th>
<th>Certainty</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Disagree</th>
<th>Least</th>
<th>Most</th>
</tr>
</thead>
</table>

18. other

<table>
<thead>
<tr>
<th>Agree</th>
<th>Certainty</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Disagree</th>
<th>Least</th>
<th>Most</th>
</tr>
</thead>
</table>
PART TWO — Demographic Information

This part of the instrument provides the researchers with demographic information necessary to describe the results of the study. When the instrument has been completed and returned, we will separate this section from the questionnaire to ensure confidentiality. This is the only time you will be asked for this information.

Directions

Please answer each of the following questions by checking the appropriate answer or by filling in the appropriate space.

Code number: 1

Area of Expertise

Phone number:

Please check the one category which best describes your expertise.

(subject matter expert)
(courseware evaluator)
(instructional designer)
(classroom teacher)

Address:

Please review the address on the mailing label below and make any necessary corrections.

Dr. Kathryn Alesandrini
N306 Lindquist Center
University of Iowa
Iowa City, Iowa 52242

Thank you for completing this questionnaire.
APPENDIX H

Round Two Delphi Instrument
Instructional Characteristics Critical For Promoting Achievement In Courseware Designed For Grades 9-12

National Delphi Study

The Ohio State University
College of Education
Instructional Design and Technology
216 Ramseyer Hall
29 West Woodruff Avenue
Columbus, OH 43210
The first part of this instrument gives you an opportunity to reconsider your original positions after reading comments of fellow panelists; the second part asks you to consider additional characteristics suggested by panel members.

We need to receive your responses by Wednesday, February 27. Please feel free to call us (614) 422-4872 if you have any questions.

Directions

Each item contains several sections. Those sections are as follows:

Responses from Round 1

This shows the percent of agreement and disagreement for each item. Your original position is marked by an X.

Comments from Round 1

Here we have included your comments—which, in most cases, have been quoted directly. In situations where we have needed to paraphrase, we hope we have not altered your original meaning.

Note from the Researchers

Only some items contain notes from us. Our notes attempt to clarify confusing terms or cast the item in a new light to help further refine the panel's position.

Your Present Position

This section requires that you restate or revise your position by checking “agree” or “disagree” and circling the degree of certainty. You also may choose to support your position. Because the Delphi Technique is designed to help reach consensus, comments from those expressing a minority opinion may be especially helpful to fellow panelists.
Example

Issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

Q-1. attractiveness of packaging

Responses from Round 1

AGREE ===== 30%

DISAGREE =========== 70%

Comments from Round 1

- you can't judge a package by its cover
- attractive packaging can help teacher and student motivation

Note from the Researchers

Remember that student characteristics (e.g., readiness, cognitive style and intrinsic learner motivation) have been omitted from this study.

Your Present Position

-------------
<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>√ DISAGREE</td>
<td>LEAST</td>
</tr>
<tr>
<td></td>
<td>MOST</td>
</tr>
</tbody>
</table>

"attractiveness" varies from person to person
PART ONE — Refinement of Issues Raised in Round 1

**Issue:** Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

1. opportunities within the program to ask questions of the program

**Responses from Round 1**

AGREE = 92%

DISAGREE = 8%

**Comments from Round 1**

- depends on the type of skill being taught
- in a simulation questions should be allowed if clarification is needed
- most important for learning highest cognitive levels
- can be an enormous waste of time if only an intrinsic motivator
- better to encourage human interaction and responses
- difficult to accomplish at a meaningful level
- if nothing else, it reduces disruptive anxiety as research demonstrated on test taking long ago

**Your Present Position**

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DISAGREE</th>
<th>LEAST</th>
<th>MOST</th>
</tr>
</thead>
</table>
Issue: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

2. Overview of skills or concepts to be learned

Responses from Round 1

AGREE ============= 85%

DISAGREE == 15%

Comments from Round 1

- software which begins with such an overview probably does nothing which can not be done more cheaply by a textbook
- most important for the designer — helpful to product quality even if students never see it — should be stated in concrete terms for student and kept as brief as possible
- overview is boring
- advance organizers can be used
- simple and brief
- can be effective in documentation but may not need to be on-line
- it's unfortunate that the researcher has linked these two
- yes, to orient and motivate

Note from the Researchers

Relying upon your comments, we've chosen to divide this item (skills from concepts) to further refine your positions. In addition we have assumed that such an overview should be included in the documentation for the teacher. We now would like you to consider whether or not the overview must be present on-line for the student.
Your Present Position

(2a) on-line overview of concepts to be learned

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

| DISAGREE | LEAST | MOST |

(2b) on-line overview of skills to be learned

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

| DISAGREE | LEAST | MOST |
Issue: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

3. opportunities for problem solving

Responses from Round 1

AGREE  ===========  100%

DISAGREE 0%

Comments from Round 1

- must be used with caution for most students don’t have extensive experience in problem-solving skills — so much of the material that emphasises these skills is usable only by a few students or with heavy instructor support — can not be taught using the “sink or swim” method
- at all levels of difficulty, especially open-ended situations
- if outcomes require this

Note from the Researchers

100% agreement requires no reevaluation of this issue unless the above comments necessitate a change in your position.
**Issue** Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

4. High degree of learner control

**Responses from Round 1**

AGREE =--------------------- 85%

DISAGREE = 15%

**Comments from Round 1**

- Depends on type of skill, level of skill, and other characteristics of the learner
- Program needs management system so student can move at own pace
- Control options should be kept simple
- Students should have reasonable control for review, seeking aid, and moving forward, but this will not facilitate learning unless the student has good motivation and good study skills — these skills are best learned in a situation where the learner gains increased control as he/she acquires techniques for using this control to aid learning (rather than simply thrashing about) — students should always have the perception that they have control — at the lowest level this may be provided by allowing reviews, opportunities to rest before continuing, and avoidance of timed advance
- Control of choice of alternatives to practice, solve or elaborate but no control that would allow the learner to short circuit necessary steps within the instruction
- Good for adults — has drawbacks for 9-12 — give control over graphics
- Degree less important than appropriateness

**Note from the Researchers**

Again we've chosen to further clarify the issue by asking you to consider the type of control (control of routine study options or control of instructional content) learners in grades 9-12 should have in order to increase achievement.
### Your Present Position

(4a) learner control of routine study options (e.g., pace of movement through the lesson, help screens, amount of practice, opportunity to review, quit option)

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

| DISAGREE | LEAST | MOST |

(4b) learner control of instructional content

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

| DISAGREE | LEAST | MOST |
Issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

5. verbal imagery (e.g., metaphors, similes, analogies, descriptive phrases)

Responses from Round 1
AGREE ~~~~~~~~~~ 54%
DISAGREE ~~~~~~~~~~ 46%

Comments from Round 1
- can sometimes be helpful but educational software should not be primarily verbal
- useful method of increasing the likelihood of associations — some caution needed where imagery may result in oversimplified view
- but words should not be overused
- research supports analogies if the object or concept used as the analogy is familiar to the learner
- but most difficult to use effectively

Your Present Position

<table>
<thead>
<tr>
<th></th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGREE</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>DISAGREE</td>
<td>LEAST MOST</td>
</tr>
</tbody>
</table>
Issue: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

6. knowledge of results

Responses from Round 1

AGREE =========== 92%

DISAGREE = 8%

Comments from Round 1

- avoid providing new information with correct feedback — students often ignore feedback when they are certain of their knowledge — some delay in feedback is useful
- should be immediate
- prescriptive information rather than correct/incorrect competency
- depends on how it is given
- this does not mean it will always help — sometimes it reduces motivation — sometimes student provides it on his own — appropriate use

Note from the Researchers

We've chosen to further clarify this item by focusing on when feedback is given. A later item addresses type of feedback desirable.

Your Present Position

(6a) immediate feedback

<table>
<thead>
<tr>
<th></th>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td></td>
<td>DISAGREE</td>
<td>LEAST MOST</td>
</tr>
</tbody>
</table>

Page 10
issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

7. depth or level of questioning

Responses from Round 1

AGREE ---------------- 85%

DISAGREE --- 15%

Comments from Round 1

- much excellent software does not question at all; it leads the student to ask questions and provides ways to discover answers and answers to one's own questions
- vague, not clear
- if it refers to requiring that students process information (retrieve, relate, generalize, etc.) in order to respond, I agree
- if it means a wide range of levels and choices by the learner, i.e., flexibility
- what matters most is the open-endedness of the questions
- inserted questions at higher levels promote learning
- if you want analysis as an outcome, call for it in the process of learning

Note from the Researchers

For clarity, this item has been reworded.

Your Present Position

(7a) questions requiring recall, comprehension, analysis, synthesis, and/or evaluation

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DISAGREE</th>
<th>LEAST</th>
<th>MOST</th>
</tr>
</thead>
</table>
Issue: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

8. opportunities for practice

Responses from Round 1

AGREE  ---------------------- 100%

DISAGREE  0%

Comments from Round 1

- must take into account instructional purpose
- should occur in wide variety of settings rather than massed immediately after the initial learning
- under student control to level appropriate for later use

Note from the Researchers

100% agreement requires no reevaluation of this issue unless the above comments necessitate a change in your position.
Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

9. Organizers embedded within the text (e.g., questions, headings, sub-headings)

Responses from Round 1

AGREE ========= 61%

DISAGREE ======= 39%

Comments from Round 1

- Text? If program is mostly text it should probably be a textbook — the computer is a visual medium
- Depends on the instructional purpose of courseware — most important in the initial learning or where motivation and learning skills are low — can be overdone to the point of being offensive to highly motivated students
- Also graphic organizers
- Naturally they don't serve alone very well but this is so with all variables — it is the mix that counts

Your Present Position

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DISAGREE</th>
<th>LEAST</th>
<th>MOST</th>
</tr>
</thead>
</table>
Issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

10. Statement of goals and objectives

Responses from Round 1

AGREE = = = = = = = = = = = 92%

DISAGREE = 6%

Comments from Round 1

- so that teacher can see how to use material to promote achievement
- depends upon whether or not the student will be held accountable for the material — may be less important if not
- much of software creates "worlds" which students may explore in their own ways
- may occur after the fact with some excellent materials that are developed by empirical, cut-and-try approaches, but usually a clear statement of objectives is the most effective method of efficient production of good materials
- can't measure effectiveness unless goals and objectives are stated
- stated simply allowing for open-ended investigation and creativity
- equally important to have goals mesh with accepted pedagogy
- research supports this but must be stated simply so learner understands
- good as advance organizer to orient and motivate, otherwise not necessary to have more than broad goals for student

Note from the Researchers

Once again we have made the assumption that the goals and objectives must be available to the teacher. We would like you to consider whether or not their availability to the learner will promote achievement.

Page 14
Your Present Position

(9a) on-line statement of goals and objectives for the learner

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>DISAGREE</td>
<td>LEAST  MOST</td>
</tr>
</tbody>
</table>
ISSUE: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

11. Opportunities to elaborate (i.e., further describe or build upon ideas within the lesson)

Responses from Round 1

AGREE: 100%

DISAGREE: 0%

Comments from Round 1

- Depends on type of skill
- Aids generalization and memory
- To encourage learner to make use of his/her ability in other areas through other media and processes
- Only if appropriate to desired outcome

Note from the Researchers

100% agreement requires no reevaluation of this issue unless the above comments necessitate a change in your position.
Issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

12. opportunities to imitate models

Responses from Round 1

AGREE  =========== 92%

DISAGREE  = 8%

Comments from Round 1

- depends on the type of skill
- not clear if human role models are included — if they’re not, they should be — exploration of assumptions behind models is valuable method of learning complex relationships or elements of systems
- but not having the narrow criteria that only those models are acceptable
- depends on the stage of the learning process — models are not always helpful especially where student framework does not allow easy organization of material

Your Present Position

<table>
<thead>
<tr>
<th></th>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>DISAGREE</td>
<td>LEAST</td>
<td>MOST</td>
</tr>
</tbody>
</table>

Page 17
ISSUE: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

13. corrective feedback
   (i.e., explanation of incorrect responses)

Responses from Round 1

AGREE  ------------------  100%

DISAGREE  0%

Comments from Round 1

- whenever possible, feedback should be intrinsic rather than right or wrong - e.g., "If that were true, this would result...(display of clearly incongruous outcome)
- during initial learning corrective feedback is important — during later practice and application better learning results if corrective feedback is withheld sometimes
- it must, however, allow for creativity and alternate responses (in a positive way)
- depends on how well it is done — it does not need to be done verbally — graphics may be better
- must be available but not necessarily regularly provided

Note from the Researchers

100% agreement requires no reevaluation of this issue unless the above comments necessitate a change in your position.
ISSUE Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

14. use of humor

Responses from Round 1

AGREE ——— 30%

DISAGREE ———— 70%

Comments from Round 1

- recent studies support this (Held, et. al)
- whimsy perhaps but jokes from a machine tend to irritate and distract rather than motivate
- must be used with extreme caution
- 21 -81% seem to be offended
- must never be directed at the learner, must never be overdone
- almost impossible to match humor to all learners — can be distracting
- I don’t believe this emotion, individual interaction, and response can be simulated
- research on humor is mixed — too many regional and context effects
- evidence on this is split - I love it but some students are put off by it and tastes differ so

Your Present Position

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DISAGREE</th>
<th>LEAST</th>
<th>MOST</th>
</tr>
</thead>
</table>

Page 19
ISSUE Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement? 

15. opportunities for success

Responses from Round 1

AGREE -------------------------- 100%

DISAGREE 0%

Comments from Round 1

- most important for those who have had minimal success in school
- must never be provided at cost of making materials appear to be too easy
- using a wide variety of opportunities at all levels
- depends on learner
- but student control should allow selection of difficulty level or previous performance or a combination of these

Note from the Researchers

100% agreement requires no reevaluation of this issue unless the above comments necessitate a change in your position. This item had the highest composite certainty level of any item on the instrument.
Issue Which instructional characteristics of courseware
designed for grades 9-12 are critical for promoting
student achievement?

16. Increased progression of skills or concepts to be learned

Responses from Round 1

AGREE ============== 77%
DISAGREE == 8%
CONFUSED === 15%

Comments from Round 1

- confusion about the meaning of increased progression
- may be helpful to structure some skills in hierarchy
  for the designer but may be a burden for the learner
- learner must have flexibility in the choice of
  degree of difficulty
- for segments of the course but not for the entire
  instructional sequence — must allow access to learners at various levels on many different
  tracks, not just one specific progression
- very important in design — courseware should allow
  for pre- and post-testing and branching so
  instruction is individualized
- agree if it means moving from the concrete to
  abstract
- as long as it doesn't preclude student control —
  indeed, it can help to facilitate building in
  student control

Note from the Researchers

This item has been reworded for clarity.
**Your Present Position**

(16a) lessons organized from the least difficult to the most difficult learning task

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DISAGREE</th>
<th>LEAST</th>
<th>MOST</th>
</tr>
</thead>
</table>
Issue: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

17. Graphic and visual cues (e.g., color coding, diagrams, charts)

Responses from Round 1

AGREE

DISAGREE 0%

Comments from Round 1

- Most important in initial learning — in some situations instructional goal may require that cues be "faded" rapidly (i.e., where students must learn to identify structures, concepts, etc. in "messy" or "real life" environment
- This can also be an organizational aid
- Effective means of teaching
- A "picture" can be worth a thousand words
- The literature supports the use of graphic cues
- Only if appropriately designed — color and illustration per se are no help

Note from the Researchers

100% agreement requires no reevaluation of this issue unless the above comments necessitate a change in your position.
## PART TWO — Additional Items Suggested by the Panel

**Issue** Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

<table>
<thead>
<tr>
<th>18. field-tested materials which have demonstrated success with target population</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AGREE</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>DISAGREE</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>19. Opportunities for teacher to &quot;author&quot; some input</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AGREE</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>DISAGREE</strong></td>
</tr>
</tbody>
</table>
ISSUE: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

22. portrays main point visually

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>DISAGREE</td>
<td>LEAST</td>
</tr>
</tbody>
</table>

Thank you for completing this questionnaire.
APPENDIX I

Round Three Delphi Instrument
Instructional Characteristics Critical For Promoting Achievement In Courseware Designed For Grades 9-12

National Delphi Study

The Ohio State University
College of Education
Instructional Design and Technology
216 Ramseyer Hall
29 West Woodruff Avenue
Columbus, OH 43210
This final instrument gives you an opportunity to reconsider your positions after reading comments of fellow panelists and to comment on your experience as a panelist. Because the panel has reached consensus on some items, the following characteristics do not require reevaluation in this instrument:

1. opportunities for problem solving
2. opportunities for practice
3. opportunities to elaborate
4. corrective feedback
5. immediate feedback
6. opportunities for success

It is very important that we receive your responses by Wednesday, March 20. Please feel free to call us collect (614) 252-0531 if you have any questions or problems.
Directions

Each item contains several sections. Those sections are as follows:

Responses from Round 2

This shows the percent of agreement and disagreement for each item. Your position from round 2 is marked by an X.

Comments from Round 2

Here we have included your comments - which, in most cases, have been quoted directly. In situations where we needed to paraphrase, we hope we have not altered your original meaning.

Note from the Researchers

Only some items contain notes from us. Our notes attempt to clarify confusing terms or cast the item in a new light to help further refine the panel's position.

Your Present Position

This section requires that you restate or revise your position by checking "agree" or "disagree" and circling the degree of certainty. You also may choose to support your position.
Example

Issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

Q-1. attractiveness of packaging

Responses from Round 2

AGREE ———— 30%
DISAGREE ———— 70%

Comments from Round 2

- you can't judge a package by its cover
- attractive packaging can help teacher and student motivation

Note from the Researchers

Remember that student characteristics (e.g., readiness, cognitive style and intrinsic learner motivation) have been omitted from this study.

Your Present Position

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>DISAGREE</td>
<td>LEAST MOST</td>
</tr>
</tbody>
</table>
REFINEMENT OF ISSUES

Issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

1. opportunities within the program to ask questions of the program

Responses from Round 2

AGREE ------------------ 100%

DISAGREE 0%

Comments from Round 2

- free-form verbal questions are difficult for program to interpret — the interactivity aspect is good but I'm not certain that asking verbal questions will promote achievement
- I take this to mean opportunities to ask "what would happen if..." and to learn by trial and error
- depends on skill being taught and on type of courseware, e.g., a drill may not require it

Note from the Researchers

This item means opportunities to query the system for help and/or additional information. We were not limiting the comment to one particular instructional strategy.
Issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

2. on-line overview of concepts to be learned

Responses from Round 2

AGREE =--------------- 92%

DISAGREE = 8%

Comments from Round 2

- better for adult learners -- research supports use of graphic organizers and overviews
- we cannot assume a teacher will always be making decisions about software -- these should be included so any learner can use them or decide to bypass them
- does a game of chess or a session at the easel begin with an overview of concepts to be learned? yet learning does occur in such experiences -- when educational software is joyful and adventuresome, it has less need for these overviews
- brief
- should be stated in concrete terms and kept as brief as possible
- studies by Hager have proven this -- make it an option that learners can access

Note from the Researchers

This issue has been rewritten for clarification.
**Your Present Position**

(2a) on-line option for overview of concepts to be learned

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>DISAGREE</td>
<td>LEAST</td>
</tr>
<tr>
<td></td>
<td>MOST</td>
</tr>
</tbody>
</table>
Issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

3. on-line overview of skills to be learned

Responses from Round 2

AGREE =========== 62%

DISAGREE ========= 38%

Comments from Round 2

- it may be difficult to provide a verbal overview of skill the learner can benefit from
- we cannot assume a teacher will always be making decisions about software -- these should be included so any learner can use them or decide to bypass them
- could be in documentation
- should be stated in concrete terms for student and kept as brief as possible
- studies by Hager have proven this -- make it an option that learner can access -- the wonderful thing about computers is that the learner can control so many aspects of the learning process
- just keep them brief -- complete documentation should be provided in the text document

Note from the Researchers

This item has been rewritten for clarification.

Your Present Position

(3a) on-line option for overview of skills to be learned

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>DISAGREE</td>
<td>LEAST  MOST</td>
</tr>
</tbody>
</table>

Page 7
Issue: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

4. Learner control of routine study options (e.g., pace of movement through the lesson, help screens, amount of practice, opportunity to review, quit option)

Responses from Round 2

AGREE: 100%

DISAGREE: 0%

Comments from Round 2

- Would still disagree if "high level of learner control" is part of the statement

Note from the Researchers

"High level" is no longer part of the statement. 100% agreement requires no reevaluation of this issue unless the above comment necessitates a change in your position.
### Issue: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

5. learner control of instructional content

#### Responses from Round 2

<table>
<thead>
<tr>
<th></th>
<th>AGREE</th>
<th>DISAGREE</th>
<th>CONFUSED</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>54%</td>
<td>39%</td>
<td>8%</td>
</tr>
</tbody>
</table>

#### Comments from Round 2

- Ideally, learner should select examples or activities that are most interesting while still attaining objectives — necessitates creation of lesson with multiple paths to objectives.
- I believe this is critical for some kinds of learning — see "Rocky's Boots" as an example of a program which is effective partly because it allows the learner to skip parts of the instructional process, and then gently steers him/her back when she/he finds that further progress is impossible without those parts.
- If in reference to "mainstream" skills — appropriate only if student has gained study skills needed for independent work.
- When appropriate — sometimes the skill or content being learned requires that the learner not be able to move ahead.
- Not necessarily 100% control.
- (I'm) confused.

#### Note from the Researchers

Assume that by controlling the content, students (whether or not competent) can bypass parts of the content.
<table>
<thead>
<tr>
<th>Your Present Position</th>
<th>AGREE</th>
<th>CERTAINTY</th>
<th>DISAGREE</th>
<th>LEAST</th>
<th>MOST</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2 3 4 5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Page 10
Issue: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

6. Verbal imagery (e.g., metaphors, similes, analogies, descriptive phrases)

Responses from Round 2

AGREE ========= 54%

DISAGREE ======= 46%

Comments from Round 2

- For learning new concepts, learner will benefit if program presents analogy to concepts familiar to learner -- use of concrete words and phrases with high imagery value can facilitate learning in early stages of new concept learning -- according to certain recent "theories" of learning, the learner can only learn new concepts via analogy to familiar ideas and prior knowledge
- Too dependent on specific environment that student has been exposed to
- The "disagree" reasons (stated in round 1) would probably hold for all characteristics surveyed in the study -- it is always a matter of appropriate use -- analogies help to mediate learning of theory but must be tied to experience, etc.
- Adventure games are mostly made of verbal imagery and they are highly motivating
- This assumes too constricted a view of courseware design options

Note from the Researchers

An issue can be "critical" without being exclusive.
<table>
<thead>
<tr>
<th>Agree</th>
<th>Certainty</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Issue: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

7. questions requiring recall, comprehension, analysis, synthesis, and/or evaluation

Responses from Round 2

| AGREE | = = = = = = = = = = 85% |
| DISAGREE | = = 15 % |

Comments from Round 2

- much excellent software does not question at all; it leads the student to ask questions and provides ways to discover answers to one's own questions

Note from the Researchers

Questions need not be explicit but may be implicit, as when the cursor appears without a stated question.

Your Present Position

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>DISAGREE</td>
<td>LEAST</td>
</tr>
</tbody>
</table>
Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

8. Organizers embedded within the text (e.g., questions, headings, sub-headings)

Responses from Round 2

AGREE ========= 62%
DISAGREE ====== 38%

Comments from Round 2

- the literature on typographical aids, graphic organizers and inserted questions support these characteristics -- maybe this should be broken down into separate question from organizers?
- should be graphic, visual, easy to spot and use not integrated with text
- add graphic representation of theory, etc., to list of examples
- especially important for students with less academic ability
- all cognitive research suggests this
- "critical" is the word which triggered my decision to disagree
- some sort of organizer is almost essential; however, its form may vary according to design

Note from the Researchers

Focus here on textual organizers. Graphic cues will be addressed later.

Your Present Position

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>DISAGREE</td>
<td>LEAST</td>
</tr>
</tbody>
</table>

Page 14
Issue: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

9. on-line statement of goals and objectives

Responses from Round 2

AGREE ========= 85%
DISAGREE === 15%

Comments from Round 2

- the learner must know what to do -- this does not always require a strict behavioral statement --
  also, this depends if the software is designed for training or educational experience in which a process is being taught -- "Rocky's Boots" is a good example of the latter
- doesn't have to be on-line
- an effective way to focus student effort if done clearly
- yes, for tutorials and drill and practice -- no, for problem solving environments such as microworlds and simulations
- how is this different from an overview of skills or concepts to be learned?
- the best software packages I've seen simply provide fascinating microworlds in which to adventure, explore and learn by trial and error; they motivate by stimulating the students' natural curiosity -- statements of goals and objectives may be perceived by many students as subtle threats of tests to come -- while they (on-line statements) may add to learning, they can also take away from the joy of learning -- good software can, in some cases, make them unnecessary by using the joy of learning as a strong intrinsic motivator
- once again, the teacher may not necessarily be involved in providing the software or using it -- the learner must have that option on-line
- inferred goals are often at least as effective as the designer's stated or intended goals
Note from the Researchers

The overview of skills and concepts is an overview of content; the statement of goals and objectives explains the behaviors expected of the learner at the end of the lesson.

<table>
<thead>
<tr>
<th></th>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>DISAGREE</td>
<td>LEAST</td>
<td>MOST</td>
</tr>
</tbody>
</table>
Issue: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

10. opportunities to imitate models

Responses from Round 2

AGREE = 92%

DISAGREE = 8%

Comments from Round 2

- I'm suspicious of what is meant by "models" and "imitate"
- realistic models may not aid initial learning of new concepts — analogies may be required first

Note from the Researchers

Model imitation often occurs when teaching problem-solving skills. For example, a lesson might show the steps involved in troubleshooting a faulty lawnmower engine and then ask the student to "imitate" that procedure in another situation.

Your Present Position

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DISAGREE</th>
<th>LEAST</th>
<th>MOST</th>
</tr>
</thead>
</table>
Issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

11. use of humor

Responses from Round 2

AGREE --- 15%

DISAGREE -------------- 85%

Comments from Round 2

- humor can help but it is difficult to define -- wit, catchy graphic effects, etc. can add much -- poor jokes, puns or regional humor can be bad
- if used with caution:
  -- never directed at the learner or groups of people
  -- never unrelated to the task (exception -- as part of a break between tasks)
  -- never as a required part of learning ("throw-away that can be ignored")
  -- should not require any but the most minimal time and effort on the part of the learner
    (avoid riddles or other forms requiring student interaction; avoid time-consuming sight gags)
- this point is not critical -- tough to use -- a broader category might get more "agree" ratings
- educational software must depend to some extent upon repetition -- very few things are funny second time around
- human emotions can't be simulated -- we are persons; the computer is a machine
- a light approach is probably appreciated but "jokes" can fall flat
Note from the Researchers

It is possible to reach a negative consensus.

Your Present Position

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>DISAGREE</td>
<td>LEAST</td>
</tr>
</tbody>
</table>
Issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

12. Lessons organized from the least difficult to the most difficult learning task

Responses from Round 2

AGREE =--------------- 85%

DISAGREE == 15%

Comments from Round 2

- however, learners should be given the option to jump to difficult tasks or choose the level of the task
- depends on the type of courseware, skills
- the generalization may have to be ignored in cases where motivation will be increased by initial attention to the most complex form of the task (e.g., where that is the form seen in "real-life")
- well-structured difficult-to-simple sequences seem to work best in early learning of everyday skills by youngest students who balk at clearly simplified (and clearly unrealistic) versions of "real problems"
- provided that students have the option to branch to any level
- would like this to read "opportunity for use of lessons in a sequence" — this one must be subject to teacher and student control
- let the student find out what he/she can and can not do and make it easy to find and use those parts of the program which can remedy specific deficits in understanding or skill
- difficulty is a relative term — learner should be equally prepared for tasks throughout the lesson — although content may increase in complexity, it may be no more difficult for the learner because the learner is equally prepared to handle learning tasks throughout the lesson — if tasks toward the end of the lesson are more difficult then the lesson has not adequately prepared the learner for the task
- this may waste time for gifted learners but there should be a recommended sequence which is least to most difficult — the learner should have the option to jump around
- it would be beneficial if the program would branch to the appropriate student level through some type of "pre-assessment" program
Note from the Researchers

For clarity, this item has been reworded.

Your Present Position

(12a) lessons should be organized for each learner from the least difficult to the most difficult learning task by some mechanism (e.g., a pre-assessment or internal self-adjustment) within the lesson

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DISAGREE</th>
<th>LEAST</th>
<th>MOST</th>
</tr>
</thead>
</table>
Issue: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

13. graphic and visual cues (e.g., color coding, diagrams, charts)

Responses from Round 2

AGREE: 92%

DISAGREE: 8%

Comments from Round 2

- I no longer feel it is critical -- important, maybe

Note from the Researchers

We no longer have consensus on this item. Please reevaluate it.

Your Present Position

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>DISAGREE</td>
<td>LEAST MOST</td>
</tr>
</tbody>
</table>
Issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

14. Field-tested materials which have demonstrated success with target population

Responses from Round 2

AGREE = 84%
DISAGREE = 8%
CONFUSED = 8%

Comments from Round 2

- This is not really an instructional characteristic.
- This (demonstrated success) is tautological. If it works then it will promote achievement.
- It's always nice to know some specifics before you decide to use materials.
- This is an important criterion for evaluation, not a characteristic which makes software good.
- Most you can do is specific population for which it was successful.
- The user must decide if it (the software) is "valid" for him/her.
- Some materials could be relevant without being field-tested, but testing would be a good guide for purchasing.
- Obviously it should be effective for instruction, but I think there's a problem when you try to rigidify implementation.

Note from the Researcher

Is this item a development procedure rather than an instructional characteristic?
<table>
<thead>
<tr>
<th></th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGREE</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>DISAGREE</td>
<td>LEAST</td>
</tr>
</tbody>
</table>
Issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

15. opportunities for teacher to "author" input

Responses from Round 2

<table>
<thead>
<tr>
<th>AGREE</th>
<th>DISAGREE</th>
</tr>
</thead>
<tbody>
<tr>
<td>------</td>
<td>----------</td>
</tr>
<tr>
<td>77%</td>
<td>23%</td>
</tr>
</tbody>
</table>

Comments from Round 2

- only if teacher tailors lesson to specific population of learners
- the best instruction takes place when an environment matches specific need of individuals
- we (teachers) don't have the time to do it well — better idea: software should provide opportunities for students to modify and extend the program
- greatly aids local acceptance but not directly necessary for student achievement if teacher is comfortable with the material and sufficiently familiar with it to provide appropriate integration with other instruction
- it should be configured with an option for a learner's own curricular setting, such as with databases
- this opportunity allows teacher to have some "control" and some ownership of the learning of his/her students — allows for the "personal touch"

Your Present Position

<table>
<thead>
<tr>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DISAGREE</th>
<th>LEAST</th>
<th>MOST</th>
</tr>
</thead>
</table>
Issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

16. presentation of real-world circumstances, not contrived information

Responses from Round 2

AGREE ========= 61%

DISAGREE ======= 31%

CONFUSED = 8%

Comments from Round 2

- this implies that only realistic examples are given -- analogies are also helpful as are abstract models, flowcharts, graphs, etc. -- if only real examples are used, learner can take information too literally and not get the abstract generalization or concept
- this leads to generalization and applications beyond the courseware
- this statement, if applied to printed materials, would deny the validity and value of all fiction -- why should it be any more applicable to software?
- must be contrived -- instruction is rarely "real world" -- simulations and work samples only approximate real world
- increases student interest and motivation but should be avoided unless designer truly has an understanding of what aspects of the "real-world" are of interest to the particular student population
- who knows the difference?
- information should be selected for its pertinence to instruction regardless of its "real" vs. "contrived" source
Your Present Position

<table>
<thead>
<tr>
<th></th>
<th>AGREE</th>
<th>CERTAINTY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td></td>
<td>DISAGREE</td>
<td>LEAST  MOST</td>
</tr>
</tbody>
</table>
Issue Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

17. open-endedness

**Responses from Round 2**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AGREE</td>
<td>------</td>
</tr>
<tr>
<td>DISAGREE</td>
<td>------</td>
</tr>
<tr>
<td>CONFUSED</td>
<td>------</td>
</tr>
</tbody>
</table>

**Comments from Round 2**

- provides more realistic learning skills and sharpens problem-solving abilities
- however, there should be an opportunity for closure
- kids need to be free to go in directions that they feel are important to them — too much of current courseware is "correct answer" oriented
- computer not appropriate for free-form responses — use tv or human teacher
- allows for higher levels of thinking, something our high school students tend to lack

**Note from the Researchers**

The original suggestion read "open-endedness and opportunities for genuine, if simulated, problem solving." We deleted the problem-solving component because the panel had already reached consensus on that characteristic. We hope we have not inadvertently altered the meaning of the original suggestion.
Your Present Position

<table>
<thead>
<tr>
<th>Agree</th>
<th>Certainty</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>Disagree</td>
<td>Least  Most</td>
</tr>
</tbody>
</table>
Issue: Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

18. portrays main point visually

Responses from Round 2

<table>
<thead>
<tr>
<th></th>
<th>AGREE</th>
<th>DISAGREE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>62%</td>
<td>38%</td>
</tr>
</tbody>
</table>

Comments from Round 2

- learners will remember the information if it has been portrayed visually -- using either analogies, real examples, or even abstract visuals such as flowcharts, graphs, Venn diagrams, etc. -- the visuals must be designed to correctly convey the information -- not thrown in for decoration -- a number of studies support the use of relevant graphics to convey main points
- many learners who choose to use courseware are visual learners -- it makes sense -- definitely reinforces
- if the main point is not presented visually and interactively, it can probably be presented more efficiently by a textbook
- a useful technique but how "critical" is it?
- I think this is rather insensitive to sightless learners
- does this "main point" mean setting the purpose? if so, definitely
Note from the Researcher

Throughout the study we are focusing on courseware for the nonhandicapped learner.

<table>
<thead>
<tr>
<th>Your Present Position</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AGREE</strong></td>
</tr>
<tr>
<td><strong>CERTAINTY</strong></td>
</tr>
<tr>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td><strong>DISAGREE</strong></td>
</tr>
<tr>
<td><strong>LEAST</strong></td>
</tr>
<tr>
<td><strong>MOST</strong></td>
</tr>
</tbody>
</table>
OPTIONAL: In the space below, please comment about your experience as a panelist and/or implications for further research.

Thank you for completing this questionnaire and for participating in this study. A final report will be forthcoming.
APPENDIX J

Results of Round Three
Results of Round Three

Round One Consensus

1. opportunities for problem solving
2. opportunities for practice
3. opportunities to elaborate
4. corrective feedback
5. immediate feedback
6. opportunities for success

Round Two Consensus

7. learner control of routine study options (e.g., pace of movement through the lesson, help screens, amount of practice, opportunity to review, quit option)
8. opportunities within the program to ask questions of the program
9. immediate feedback

Round Three Consensus

No additional consensus.
FINAL REFINEMENT OF ISSUES

**Issue** Which instructional characteristics of courseware designed for grades 9-12 are critical for promoting student achievement?

1. opportunities within the program to ask questions of the program

   **Responses from Round 2**
   
   AGREE ============== 100%
   
   DISAGREE 0%

   **Comments from Round 3**
   
   None.

2. on-line option for overview of concepts to be learned

   **Responses from Round 3**
   
   AGREE ============== 85%
   
   DISAGREE === 15%

   **Comments from Round 3**
   
   - some very effective software lacks such an option; therefore, it must not be crucial
   - data show that such guides improve efficiency of instruction but I would not say this is actually crucial. The same effect can be produced in interactive materials by careful choice of directions and feedback for embedded challenge items
3. on-line option for overview of skills to be learned

Responses from Round 3

AGREE =========== 85%

DISAGREE === 15%

Comments from Round 3

- not as important as the choice of embedded practice situations though

4. learner control of routine study options
   (e.g., pace of movement through the lesson, help screens, amount of practice, opportunity to review, quit option)

Responses from Round 3

AGREE =========== 100%

DISAGREE 0%

Comments from Round 3

None.

5. learner control of instructional content

Responses from Round 3

AGREE =========== 46%

DISAGREE =========== 54%

Comments from Round 3

- not critical but very often desirable
- if the learner control could evoke a "necessary" (design goal) part of the instruction completely (not simply delay it), then the learner control is antithetical to achievement. Unless carefully used with full awareness of student capabilities, learner control can lead to grossly inefficient learning and even to learner frustration. At the same time, learner control of education should be a major goal of 9 - 12 education. Many proponents of learner control think only of this goal, however, and fail to understand the distinction between prepartation
for independent learning (mostly non-school or post-secondary) and school-supported learning (which is obligated to provide efficient learning of basics at low cost). Learner control is thus a necessary part of some 9-12 instruction but not critical (or even desirable) in all. - still very much depends on the learner, content, grade level, etc.

6. verbal imagery
   (e.g., metaphors, similes, analogies, descriptive phrases)

Responses from Round 3

AGREE =========== 54%
DISAGREE ========= 46%

Comments from Round 3

- some excellent and effective pieces of software are entirely non-verbal

7. questions requiring recall, comprehension, analysis, synthesis, and/or evaluation

Responses from Round 3

AGREE =============== 92%
DISAGREE == 8%

Comments from Round 3

- respondent wanted item reworded to take the emphasis off text
8. organizers embedded within the text 
(e.g., questions, headings, sub-headings)

Responses from Round 3

AGREE ========= 62%

DISAGREE ====== 38%

Comments from Round 3
- good courseware should be primarily non-textual
- depends on skills
- text (and textual organizers) aren't essential

9. on-line statement of goals and objectives

Responses from Round 3

AGREE ============= 85%

DISAGREE == 15%

Comments from Round 3
- doesn't have to be on-line

10. opportunities to imitate models

Responses from Round 3

AGREE ============== 92%

DISAGREE = 8%

Comments from Round 3
- however, it is also useful to learn that such imitation must be followed only when appropriate (or that the model may be imperfect)
- depends on the type of courseware
11. use of humor

Responses from Round 3

AGREE == 8%

DISAGREE =========== 92%

Comments from Round 3

- in some subject areas use of humor has been shown to lead to significantly better retention (e.g., absurd outcomes generated by commonly encountered erroneous interpretations or applications are demonstrated in the instruction) -- students report very high recall for these situations and show substantially better avoidance of the specific classes of errors that were demonstrated in humorous fashion

- believes style a critical concern -- educational films are often criticized as being dull or all the same

12. lessons organized from the least difficult to the most difficult learning task

Responses from Round 3

AGREE =========== 85%

DISAGREE === 15%

Comments from Round 3

- in some cases it may be more efficient to let the learner try whatever level s/he wishes and then provide help with finding the prerequisites if s/he runs into difficulty

- still depends on the type of courseware, learners, skill
13. graphic and visual cues
(e.g., color coding, diagrams, charts)

Responses from Round 3

AGREE ============= 85%
DISAGREE === 15%

Comments from Round 3

- depends on the learner and skills being taught
- such cues are useful — Color is attractive and preferred by most students and instructors, but the data almost never show significant advantages for heavy graphic cues — I also shift to disagree
- it depends on the type of program — In science, problem solving etc. can be extremely important — in a drill and practice it may be superfluous
- if these are not used, perhaps the material could be presented more cost-effectively in printed form

14. field-tested materials which have demonstrated success with target population

Responses from Round 3

Impossible to evaluate in the traditional manner due to researcher error

Comments from Round 3

- a development procedure rather than an instructional characteristic
- a development procedure
- not an instructional characteristic — believes the critical criterion is that "we know what it (the lesson) does" i.e., that the user knows that certain interpretations as to what happens to a known population are more plausible by data & a line of argument" — then user can decide whether or not to try it out — it is essential for a developer to take responsibility and to gather appropriate data and bring out the kind of argument suggested
- the fact that the lesson has been field-tested with a particular population helps the consumer make the decision to incorporate that software in a student's instructional plan
- it is a development procedure but it is important
- experienced designers can do an amazing job of
  matching student needs without field-test revisions
  -- however, most have gained this ability by doing
  a lot of field testing in the past -- for designers
  without an intimate knowledge of the design student
  population, the subject matter, and techniques of
  individualized instruction; field testing is a
critical requirement for effectiveness
- (I believe it is a design rather than an
  instructional characteristic) it is,
  however, the single most important thing to feature
to look for when purchasing courseware

15. opportunities for teacher to "author" some input

Responses from Round 3

AGREE =========== 69%

DISAGREE ======= 31%

Comments from Round 3

- (no,) it takes the control of design and
  replicability of results out of the designer's
  hands
- allowing students to recode the program (as
  suggested earlier) presents some real hazards
- certainly not for all software but I find teachers
  who want the flexibility to tailor programs to
  students' needs -- the option should be available in
  some instances
- I think students should be able to author as well
- this is no more critical with computer software
  than it is with textbooks
- not sure this isn't a design feature -- it is a
curriculum feature that impacts actual behavioral
language
16. presentation of real-world circumstances, not contrived information

Responses from Round 3

AGREE ========= 46%

DISAGREE ========= 56%

Comments from Round 3

- all we can do is develop a "work model" — see Bunderson et al. Instructional Science, 10(81), 205-215.
- what is needed is a good mix of real world and contrived
- depends on the instructional strategy and content

17. open-endedness

Responses from Round 3

AGREE ========= 46%

DISAGREE ========= 46%

CONFUSED == 8%

Comments from Round 3

- open-endedness may very well contribute to some kind of student achievement but not to achievement of all objectives — for example, I am not sure how open-endedness could contribute to teaching ten new vocabulary words in chemistry — until artificial intelligence becomes more real, how could a computer respond to open-ended answers
- too vague
- still think this is better accomplished with the human touch
- this may exceed the current technology for many topic areas
18. portrays main point visually

Responses from Round 3

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AGREE</td>
<td>=======</td>
<td>62%</td>
</tr>
<tr>
<td>DISAGREE</td>
<td>======</td>
<td>30%</td>
</tr>
<tr>
<td>CONFUSED</td>
<td>==</td>
<td>8%</td>
</tr>
</tbody>
</table>

Comments from Round 3

- depends on type of courseware
- important in some situations but not critical in all
- too vague

OPTIONAL: In the space below, please comment about your experience as a panelist and/or implications for further research.

1. appreciated revisions made on input -- variety of comments suggests a complexity not accessible by Delphi -- perhaps "critical incident technique" would be a good follow-up -- thank you for the opportunity to participate & best wishes

2. thank you for the opportunity to express opinions -- researchers accurately reflected my comments and concerns -- many characteristics were so general to achieve consensus that I wonder how useful they will be

3. I enjoyed participating in the study. It was interesting to ponder important vs. critical and to be swayed by the panelists comments. Now that we have consensus on what characteristics promote student achievement, we need to focus on components needed to truly integrate computers in 9-12 instruction.

4. Besides degree of certainty, degree of importance should also be asked. I found that some items were more critical than others -- not all features applied to all courseware

5. From the comments of the panelists, I would infer that they had a broad range of experiences. Some obviously had little idea of what the questions were getting at, therefore, I would carefully weigh their responses against those of more experienced respondents.
Also it is difficult to say that one single group of features affects all learning outcomes in the same way. For example, software that uses drill is directed at a different cognitive outcome than software that is directed at problem solving or creativity. Therefore, they must be judged against different criteria. The guidelines I have included reflect an attempt to set forth such an evaluation tool.

6. Thanks. I enjoyed it.

7. An on-line Delphi using one of the national computer networks would have been interesting. I sensed from some comments that other participants had specific information that I would like to have known more about before making a decision. Others made comments that seemed contrary to my experience and I would have liked to have made certain that these people were aware of this information.

My overall feeling with the current method is that I did not modify any of my views, but rather expended most of my effort in understanding the questions. I hope that my answers were to the same questions answered by most others and not to my own idiosyncratic interpretation of the question.

I enjoyed participating in the study. Best wishes in your final analysis of the data.

8. I suspect that your results will be questionable for two reasons:

1) In several instances, disagreement seemed to result from widely divergent interpretations of your words. Some of these ambiguities remain, even in the third round.

2) The method by which participants were selected leaves room for doubt that we are highly qualified to speak on this subject.

Participation was an interesting experience; I look forward to reading the summary report. By the way, the Droste chocolate bar was a most welcome gift.

9. Like many questionnaires, this one still operates at a high level of generality. I suspect that many conclusions (including my own) would not hold up when applied to specific programs in an actual teaching context.

10. [Synopsis of letter] enjoyed participating — believes serious error to try to come up with one set of evaluation criteria for all courseware — only criterion s/he has found effective (other than content and technical accuracy) is evidence of testing with students
-- feels we're wasting money on large evaluation schemes - need to concentrate instead on cataloging accurate courseware and letting individuals evaluate it themselves -- changed some of her/his answers on last two rounds because felt degree of agree/disagree not being factored in

11. The experience causes me to want to meet the other panelists to discuss the issues we have identified.

The real challenge is to convince software creators to follow our suggestions. It would also be helpful to share examples of current courseware that meet the criteria.