INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the original text directly from the copy submitted. Thus, some dissertation copies are in typewriter face, while others may be from a computer printer.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyrighted material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is available as one exposure on a standard 35 mm slide or as a 17" × 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. 35 mm slides or 6" × 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
The identification of evaluation criteria items and item importance of drill/practice and educational game software programs

Jeng, Yoau-Chau, Ph.D.

Iowa State University, 1988
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 \( \checkmark \).

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 \( \checkmark \)
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 \( \checkmark \)
16. Other ________________________________________________
    ________________________________________________
    ________________________________________________

UMI
The identification of evaluation criteria items and item importance of drill/practice and educational game software programs

by

Yoau-Chau Jeng

A Dissertation Submitted to the Graduate Faculty in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major: Industrial Education and Technology

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa
1988

Copyright (c) Yoau-Chau Jeng, 1988. All rights reserved.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER I. INTRODUCTION</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statement of the Problem</td>
<td>5</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>6</td>
</tr>
<tr>
<td>Research Questions</td>
<td>6</td>
</tr>
<tr>
<td>Hypotheses of the Study</td>
<td>7</td>
</tr>
<tr>
<td>Assumptions of the Study</td>
<td>9</td>
</tr>
<tr>
<td>Limitations of the Study</td>
<td>9</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>10</td>
</tr>
<tr>
<td>Procedure of the Study</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER II. REVIEW OF THE LITERATURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview</td>
<td>15</td>
</tr>
<tr>
<td>Quantitative vs Qualitative Methods</td>
<td>19</td>
</tr>
<tr>
<td>Evaluation Criteria</td>
<td>22</td>
</tr>
<tr>
<td>Weighting Criteria</td>
<td>40</td>
</tr>
<tr>
<td>Summary</td>
<td>42</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER III. METHODOLOGY</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population and Samples</td>
<td>45</td>
</tr>
<tr>
<td>Instrument Development</td>
<td>45</td>
</tr>
<tr>
<td>Data Collection</td>
<td>47</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>47</td>
</tr>
<tr>
<td>Descriptive analysis</td>
<td>48</td>
</tr>
<tr>
<td>Inferential analysis</td>
<td>48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER IV. RESULTS AND FINDINGS</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey Response</td>
<td>49</td>
</tr>
<tr>
<td>Pilot Study of the Data Collection Instrument</td>
<td>50</td>
</tr>
<tr>
<td>Response of the Basic Data</td>
<td>51</td>
</tr>
<tr>
<td>Results of the Hypothesis Testing</td>
<td>55</td>
</tr>
<tr>
<td>Hypothesis 1</td>
<td>55</td>
</tr>
<tr>
<td>Hypothesis 2</td>
<td>56</td>
</tr>
<tr>
<td>Hypothesis 3</td>
<td>58</td>
</tr>
<tr>
<td>Hypothesis 4</td>
<td>60</td>
</tr>
<tr>
<td>Hypothesis 5</td>
<td>60</td>
</tr>
<tr>
<td>Hypothesis 6</td>
<td>60</td>
</tr>
<tr>
<td>Hypothesis 7</td>
<td>61</td>
</tr>
<tr>
<td>Hypothesis 8</td>
<td>61</td>
</tr>
<tr>
<td>Hypothesis 9</td>
<td>61</td>
</tr>
<tr>
<td>Results of the Research Questions</td>
<td>61</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER V. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restatement of the purpose</td>
<td>83</td>
</tr>
<tr>
<td>Restatement of the research questions</td>
<td>83</td>
</tr>
<tr>
<td>Conclusions</td>
<td>86</td>
</tr>
<tr>
<td>Hypothesis 1</td>
<td>86</td>
</tr>
<tr>
<td>Conclusion of Hypothesis 1</td>
<td>86</td>
</tr>
</tbody>
</table>
LIST OF TABLES

TABLE 1. The number and percent of respondents in sample ........................................ 52
TABLE 2. The sex of respondents ........................................ 52
TABLE 3. The number and percent of school teachers classified by teaching level ............ 52
TABLE 4. Types of microcomputer software programs used in classroom teaching .......... 53
TABLE 5. The number of years of teaching experience ........................................ 54
TABLE 6. The number of years experience with computers ........................................ 54
TABLE 7. The number and percent of teachers with experience in using drill/practice software programs ........................................ 55
TABLE 8. The number and percent of teachers with experience in using educational software programs ........................................ 55
TABLE 9. Situation in which computer software is most useful ........................................ 56
TABLE 10. The split-plot ANOVA of teacher groups and types of software programs for content quality evaluation criteria ........................................ 57
TABLE 11. The split-plot ANOVA of teacher groups and types of software programs for instructional quality evaluation criteria ........................................ 58
TABLE 12. The split-plot ANOVA of teacher groups and types of software programs for technical quality evaluation criteria ........................................ 59
TABLE 13. The item importance priority, mean, and standard deviation of evaluation criteria
TABLE 14. The item importance priority, mean, and standard deviation of evaluation criteria from combined-group for drill/practice software programs ............... 62

TABLE 15. The item importance priority, mean, and standard deviation of evaluation criteria from elementary school teachers for drill/practice software programs ............. 68

TABLE 16. The item importance priority, mean, and standard deviation of evaluation criteria from elementary school teachers for educational game software programs ........... 71

TABLE 17. The item importance priority, mean, and standard deviation of evaluation criteria from secondary school teachers for drill/practice software programs ............. 74

TABLE 18. The item importance priority, mean, and standard deviation of evaluation criteria from secondary school teachers for educational game software programs ........... 77

TABLE 19. The frequency of the average score range of evaluation items for drill/practice software programs ............... 80

TABLE 20. The frequency of the average score range of evaluation items for educational game software programs ............... 80

TABLE 21. Cronbach's alpha reliability coefficient for teachers and evaluation criteria qualities .................. 81

TABLE 22. Spearman rank correlations based on item means between teacher groups and software programs ............... 81
CHAPTER I. INTRODUCTION

Everyone today it seems has an interest in microcomputers—parents who are pressuring the schools to buy them lest their children be left behind, teachers who want to use them not only for teaching skills but to manage instruction and handle other clerical details of their jobs, administrators who wish to manage all of their administrative tasks. Many people refer to this new technology as the microrevolution and compare its potential impact to that of the industrial evolution (Mathews & Winkle, as cited in Truett, 1984).

Microcomputer, with its simplicity and versatility, coupled with relatively lower cost, greatly expanded the use of computers in education. According to Curriculum Information Center's survey (1984-1985) of microcomputers in schools of the United States, the 1983-1984 school year was a banner year for the installation of microcomputers to assist instruction. The number of micros in public schools jumped by 75% from Fall 1983 to Fall 1984. This represents a phenomenal growth rate considering the fact that micros were already installed in 68% of public schools in the Fall of 1983. By Fall 1984, over 630,000 microcomputers were available for student instruction in public, private, and Catholic schools. At the present rate of increase in
installations, the next survey (1985-1986) did show as many as 1,000,000 microcomputers in our schools, representing a cumulative investment of about one billion dollars for equipment, plus countless other expenditures for software, supplies, peripherals, and training.

At the same time, there has been a proliferation of software developed to be used with microcomputers in elementary and secondary schools for the past several years. Software experts have estimated that by 1985 over 7,700 educational software packages were available in the United States and Canada, with up to 2,000 packages being added each year by more than 700 educational software producers (MICROgram, 1985). According to Lathrop and Goodson (1983), only about one percent of these programs have been evaluated. Much of this software varies greatly in its quality and scope, and the best advice that can be given to educators seeking software is "vaveat emptor"—buyer beware. A report released in 1982 by the Office of Technology Assessment (Truett, 1984) of the Congressional Board of the ninety-seven Congress cited four reasons to explain the lack of quality educational software. In the first place, much of the technology is still new. It takes time to learn how to use it and the early attempts suffer from this learning process. Second, production of high quality educational
software is expensive. Some large firms that have the necessary capital to produce educational software hesitate to risk development money in a relatively new and uncertain market. Third, the programmers and curriculum experts qualified to produce educational software are in short supply. Finally, some firms cite the lack of adequate property protection (e.g., copyright, patents) for their information products as a barrier to investment in development.

Roblyer (1981) and Cohen (1983) state that the most common procedure of instructional software are cottage industry firms made up of enterprising individuals who lack expertise in systematic design but have rushed in to fill the vacuum created by the demand for educational software. Hall, Comer and Merrill (as cited in Roblyer, 1983) added that software development is the purview of anyone with a microcomputer, many of whom are ill-prepared to develop quality materials. There is a widespread interest for teachers to create their own materials for the technology, and the literature of programming and authoring offer a set of quick, simple cookbook-like procedures (Roblyer, 1983). As a result there is a proliferation of software which has doubtful instructional value.
As the use of microcomputers proliferates in educational setting, the production of suitable instructional materials for the microcomputer has become an issue of great importance and concern to educators. There is a great need to determine what exactly is "good" instructional software and what is inadequate. As more software programs become available, how to discriminate between the effective and ineffective ones is a critical problem. Consequently, the number of individuals and organizations who recommend guidelines for evaluating software seems to increase proportionately. The first problem which seems to reoccur on evaluation guidelines is that in almost all cases criteria are given equal weight in the evaluation process. That is, each evaluation criterion is given equal importance in evaluating the software. Obviously, not all features of a peace of software contribute equally to the effectiveness of the total problem. With this in mind, it would seem logical the evaluation guidelines be structured so that a composite rating putting more emphasis on a feature such as "interactiveness" than on something else like "Feedback is personalized." Evaluation criteria, therefore, need to be categorized into a hierarchy which will reflect the degree of contribution each criterion makes to the effectiveness of
the overall program. The second problem prevails quite often in evaluation guidelines is that criteria are usually designed to be applied to all software programs regardless of the teaching strategy used to deliver the content. Classroom teaching is very often evaluated without regard for the context of the objectives taught in the lesson. Few educators would argue, however, that very different skills are needed to make individual strategies such as inquiry, lecture, drill, and group discussion effective independently. It would seem to follow, then, that the same criteria applied to evaluating a drill/practice software program might be inappropriate when applying them to evaluate the effectiveness of a computer-based simulation (Caldwell, 1983).

Statement of the Problem

The problem of this study was to investigate:

1. Which evaluation criteria will be necessary for the evaluation of drill/practice and educational game software programs.

2. Importance of each criterion in the process of evaluating drill/practice and educational game software programs.
Purpose of the Study

The purpose of the study was:
1. To identify which evaluation criteria will be necessary for the drill/practice and educational game software programs.
2. To identify each evaluation criterion the average importance on evaluation guidelines in the process of evaluating drill/practice and educational game software programs.
3. To provide school teachers, administrators and individuals with reference information for the evaluation of microcomputer software programs.

Research Questions

This study was designed to answer the following questions:
1. Is there any difference between the evaluation criteria for the drill/practice and educational game software programs?
2. Does each evaluation criterion gain the same importance (weighting score) in the process of evaluating drill/practice and educational game software program?
3. If not, what will be the average importance (average scale value) for each evaluation criterion?

Hypotheses of the Study

1. There is no significant difference between the scale values of content quality evaluation criteria from responses of elementary school teachers and secondary school teachers.
   Ho: $\mu_1 = \mu_2$   Ha: $\mu_1 \neq \mu_2$

2. There is no significant difference between the scale values of instructional quality evaluation criteria from responses of elementary school teachers and secondary school teachers.
   Ho: $\mu_1 = \mu_2$   Ha: $\mu_1 \neq \mu_2$

3. There is no significant difference between the scale values of technical quality evaluation criteria from responses of elementary school teachers and secondary school teachers.
   Ho: $\mu_1 = \mu_2$   Ha: $\mu_1 \neq \mu_2$

4. There is no significant difference between the scale values of content quality evaluation criteria from drill/practice and educational game software programs.
5. There is no significant difference between the scale values of instructional quality evaluation criteria from drill/practice and educational game software programs.

\[ H_0: \mu_1 = \mu_2 \quad \text{Ha: } \mu_1 \neq \mu_2 \]

6. There is no significant difference between the scale values of technical quality evaluation criteria from drill/practice and educational game software programs.

\[ H_0: \mu_1 = \mu_2 \quad \text{Ha: } \mu_1 \neq \mu_2 \]

7. There is no interaction between the scale values of content quality evaluation criteria from school teachers and software programs.

\[ H_0: \alpha \beta = 0 \quad \text{Ha: } \alpha \beta \neq 0 \]

8. There is no interaction between the scale values of instructional quality evaluation criteria from school teachers and software programs.

\[ H_0: \alpha \beta = 0 \quad \text{Ha: } \alpha \beta \neq 0 \]

9. There is no interaction between the scale values of technical quality evaluation criteria from school teachers and software programs.

\[ H_0: \alpha \beta = 0 \quad \text{Ha: } \alpha \beta \neq 0 \]
Assumptions of the Study

This study was conducted under the following assumptions:

1. The procedure for selecting the research subjects was valid and adequate for making inferences for the general population.
2. The survey questionnaire was an appropriate measure for collecting information on evaluation criteria items and each item's importance.
3. Respondents interpreted the questionnaire items correctly.
4. Respondents responded to the questionnaire honestly.

Limitations of the Study

This study was conducted under the following limitations:

1. The evaluation criteria items and each item's importance was identified from the questionnaire survey method, not from the correlational research method, which means verifying the relationship between the evaluation criterion and its effect on the achievement of the learner.
2. Respondents were requested to determine the weighting score (importance) for each evaluation criterion is subjective, and therefore may be influenced by their experiences of using microcomputer software programs.

3. The study was based on responses from a sample of elementary and secondary school teachers in the state of Iowa. Thus, any generalization beyond this state cannot be assumed.

Definition of Terms

**Computer-assisted instruction**: An interaction between a student controlled display, and a response-entry device for the purpose of achieving educational outcome (Bunderson, 1976).

**Correlational research**: In which an attempt is made to discover or clarity relationship through the use of correlation coefficient. It tells the researcher the magnitude of the relationship between two variables A and B, but they cannot be used to determine whether A causes B, or B causes A, or whether a third variable, X, causes both A and B (Borg and Gall, 1983).

**Drill and practice**: Designed to supplement regular instruction received by providing a means by which concepts
presented and developed in the classroom can be practiced and refined at the computer (O'Neil and Richardson, 1977).

**Educational game:** A situation where students have to know certain fact, perform certain skills, or demonstrate mastery of certain concepts; winning depends upon mastery of these cognitive skills (Seidner, 1976).

**Elementary schools:** Schools with grades K-6.

**Hardware:** Computer related equipment such as disk drives, microprocessors, cables, cassette players, printers, and monitors.

**Microcomputer:** A computer that costs roughly $200 to $6,000 with the major computational capabilities concentrated in one electronic component called a "chip". Generally, it can not only do computing but can also communicate with terminals and store relatively large quantities of data. The functional components are the same as a large computer system which includes: input, memory, output, and a central processing unit (CPU). Memory usually ranges from 4K (enough space to store about 4,000 characters) to 64K (about 64,000 characters). The commonly available language is BASIC (Frederick, 1980; Milner, 1980).

**Program:** A set of instructions written for the microcomputer to carry out its intended functions (Marshall, 1983).
Secondary schools: Schools with grades 7-12.

Simulation: The dynamic execution or manipulation of a model of some object system so that the student can interact with and become part of that simulated reality (Seidner, 1976).

Software: Instructions given to the microcomputer and supported data/material necessary for the computer to accomplish it's designated tasks (Marshall, 1983).

Tutorial: Intended to stand alone as an instructional entity, teacher's rules and concepts embodied in the subject matter as well evaluating the students' comprehension of these concepts (O'Neil and Richardson, 1977).

Weighting score: The indicator which presents the relative importance of each evaluation criterion in this study.

Procedure of the Study

1. To synthesize, by reviewing previous software evaluation research literature, the evaluation criteria for the microcomputer software programs.
2. To determine the population of the study.
   According to Curriculum Information Center's School Directory, school year 1983/1984, there were 1,506 elementary and secondary schools in
the state of Iowa which using microcomputers in classroom teaching. Those schools were identified as the population of this study.

3. To select the sample from the population. A stratified (elementary schools and secondary schools) sampling technique was employed to select the sample size of 301 schools from the population.

4. To develop an evaluation criteria survey instrument (questionnaire) which contained basic information and evaluation criteria list for the drill/practice and educational game software programs.

5. To verify the content validity and the appropriateness of questionnaire item construction. Assistance was sought from the experts at Iowa State University.

6. To revise the questionnaire based on the recommendations of the experts.

7. To conduct a pilot study with people selected from the population to try out the questionnaire.

8. To finally revise the questionnaire from the results of pilot study data.

9. To conduct a field test.
10. To analyze the data and test the hypotheses: a. A descriptive analysis was used to determine the mean and standard deviation of scale value for each evaluation criterion of the drill/practice and educational game software programs. b. A split-plot analysis of variance (ANOVA) (2 x 2) was applied to determine if a statistical significance existed between the scale values of evaluation criteria from elementary school teachers and secondary school teachers. c. A split-plot analysis of variance (2 x 2) was applied to determine if a statistical significance existed between the scale value of evaluation criteria from drill/practice and educational game software programs. d. A split-plot analysis of variance (2 x 2) was applied to determine if an interaction existed between the scale values of school teachers and software programs.

11. To finish the research report based on the results of data analysis.
CHAPTER II. REVIEW OF THE LITERATURE

Overview

The use of microcomputers and educational software has been growing dramatically during the last several years. By 1985 over 7,700 educational software packages were available in the USA and Canada, with up to 2,000 packages being added each year by more than 700 educational software producers (MICROgram, 1985). Despite the rapid growth of the educational software industry, however, concerns over software quality persist.

In the early 1980s, when the development of industrial software for microcomputers was still in its infancy, there were already serious concerns about the quality of educational software (e.g., Budoff & Hutten; Hofmeister, as cited in Dudley-Marling & Owston, 1987). However, there was considerable optimism that this situation would quickly be rectified and, by 1983 Lathrop and Goodson concluded that simple drill software was being replaced by courseware that took full advantage of the capabilities of the new computers and display screens using color, detailed graphics, and sound advantageously. Similarly, Truett (1984), while conceding the poor quality of early programs, argued that the overall quality and variety of instructional software
had improved. Bitter (1984) also envisioned a bright future for educational software, asserting that sophisticated, educationally sound software was rapidly becoming available for all subject areas and grade levels. Eisele (1985) noted particularly the tremendous improvement in applications software (e.g., graphics, etc.) used in education. Torgeson (1984) was so positive about computer-assisted instruction (CAI) that it was concluded that as little as ten minutes a day using their current CAI software could significantly affect students' academic achievement (Dudley-Marling & Owston, 1987).

The consensus of opinion, however, was that, although there were some excellent programs available, the overall quality of the instructional software remained relatively poor (Allen; Bitter; Holznazel; Jenson; Olvard; Staples, as cited in Dudley-Marling & Owston, 1987).

Observers of educational software have cited a number of specific problems with programs, including poor pedagogy, amateurish programming, and inadequate documentation (Staples, 1985). The most common explanation for this situation has been that most software is written either by persons with expertise in computers but not in education, or expertise in instructional theory but not computers (Allen; Bitter; Torgesen, as cited in Dudley-Marling & Owston,
1987). Johnson (1984) cited another, more pervasive problem. Few educational materials, including educational software, are field tested with actual students prior to distribution.

There has also been concern expressed that educational software is limited in its scope and versatility to accommodate individual differences in student learning (Kolich; Hummel & Senf, as cited in Dudley-Marling & Owston, 1987). Probably the most frequent criticism of software, however, is the predominance of drill and practice programs (Allen; Bitter; Chandler; Hummel & Balcom, as cited in Dudley-Marling & Owston, 1987), which are viewed as dull and unimaginative (Cohen & Schwartz, 1983), focusing on lower order thinking skills (Holznazel, 1983), and which do not take full advantage of the capabilities of the microcomputer (Bialo & Erikson; Bitter; Hummel & Balcom, as cited in Dudley-Marling & Owston, 1987). Former U.S. Secondary of Education Terrel Bell (1984), for example, estimated that less than five percent of then available educational software took the advantage of the unique abilities of microcomputers to improve teaching and learning.

While many educators bemoan the preponderance of drill and practice software, which is based on a behavioral view of learning, others defend its use (e.g., Roblyer, 1986).
Jensen (1985) argued that the problem with much of the available CAI software was that software developers frequently overlook important behavioral principles like transfer of stimulus control and the use of adequate reinforcers.

Despite the abundance of opinions regarding the state of educational software, there have been few objective evaluations of the current pool of educational software, although the data which are available tend to support those who are critical of the state of educational software.

The Educational Products Information Exchange (EPIE) produces large numbers of software evaluations which are made available to its subscribers. EPIE staff reported that only five percent of the software they have evaluated up to 1985 had been rated as "exemplary" (MICROgram, 1985) and only about one quarter had met even minimal EPIE standards (Komoski, cited in Benderson, 1985). EPIE later confirmed these findings by applying the California State Department of Education "Guidelines for Educational Software Evaluation for California Schools" to representative sample of educational software (MICROgram, 1985). Similarly, MicroSIFT, another software evaluation service, has been able to "highly recommend" only 17 percent of the software it evaluated up to 1985 (Benderson, 1985). Further support
for these conclusions comes from a provincial software project in the Canadian province of Alberta, which in 1984 was unable to recommend nine out of ten software products previewed for use in Alberta schools (Alberta Education, 1985).

Quantitative vs Qualitative Methods

Obviously there are many ways to approach the problem of evaluating computer programs in the classroom, but in general they can be sub-divided into two broad categories based upon the nature of the data to be collected and the use to which it may be put. The first category contains all those approaches which involve measurement of some kind which may, in certain circumstances, be suitable for statistical analysis. These are called "quantitative methods". Those which do not involve measurement, but require the collection of people's views or impressions, or set out to describe how a program has been used, or the behavior of the children, while using it are basically descriptive and are called "qualitative methods" (Blease, 1986).

We can see that quantitative evaluation emphasizes a more objective analysis of the particular parts of programs or products, qualitative evaluation emphasizes a more
subjective, intuitive analysis of the totality of programs or products (Day, 1984).

Because of the more subjective nature of qualitative methods, critics wereraised by some evaluators. Roblyer (1981) believed that subjective criteria received far too much emphasis. Steffin (1983) asserted that the essential component of changes, modifications, or additions to the learner's behavior may be identified. To introduce concerns reflected in such terms as "school social purpose", motivation, or aesthetic quality, is to introduce a level of methodological ambiguity which serves to confuse efforts to remedy or improve existing instructional systems. Caldwell (1983) stated that few of the criteria used to evaluate software have been validated through research and experimentation; instead, they often find their basis in speculation and intuition only. Criteria for evaluation are frequently highly inferential in nature which of course, makes them highly subjective. Subsequently, subjectivity can only serve to lower reliability among separate ratings. Blum (1982) observed wide variance in ratings of software by three or more viewers. This variance was attributed to: (a) error in scoring due to inadequate training or background of the reviewers; and (b) subjective judgments of the evaluators on items that were highly inferential.
On the other hand, comments about quantitative methods were proposed by some evaluators. Blease (1986) stated:

"How can you control for differences in the quality of presentation to the different groups when different methods or media are being used? How do you know that the quality of presentation of a computer program—one that exploits the potential of the medium—is equivalent in quality to the presentation in a book or a lecture or a television program?" (p. 100).

Bates (1981) suggested that one of the clear findings from experimental research is that learning gains tend to vary more within than between media.

Blease also was concerned such that experiments are unlikely to be part of a normal everyday teaching program in that they do not significantly represent the situation in which the program to be evaluated would normally be used. Even if all of the above problems can be overcome, its not always easy or even possible to pinpoint or isolate what it is about a particular program that it so effective.

Holznagel (1981) commented:

"It would be desirable in most cases to base an evaluation on student use of a package.... Such evaluation might include pretests and post-tests, observation of student use, and anecdotal records from the observation.
That kind of activity, however, is expensive and time-consuming. The MicroSIFT project will not be able to implement this level of evaluation in most cases" (p. 140).

That means cost and time is another important constraint for the quantitative evaluation of software programs.

Evaluation Criteria

Evaluation criteria can be classified into two basic types: internal and external. Internal criteria refer to the intrinsic characteristics of the program or product being evaluated. External criteria refer to actual performance results (Day, 1984). Most of the software evaluation literature discusses the various internal criteria that experts feel software should have in order to be effective. Generally, evaluation researchers classify software into three broad categories: content quality, instructional quality, and technical quality (Wager, 1981). Content quality address such issues as the accuracy of concepts and ideas presented in the software, appropriateness of the difficulty levels for the intended users, and relevance of the software to the subject matter. The instructional quality criteria deal with the actual process of the instruction. Typical criteria include:
proper learner control of the pace of learning, appropriate use of computer graphics and sound, adequate learner instructions, and appropriate use of the capabilities of the microcomputer. Technical quality usually refers to the actual programming quality of the software (Does it run? Is it easy to use?) as well as the adequacy of the documentation (Day, 1984).

Cohen (1983) identified two kind of attributes that should be used in development and evaluation: (1) those that are generic to instructional design, and (2) those that are necessary to consider in the design of courseware.

Generic to instructional design:
1. Specified target audience/intended users
2. Specified learner entry competencies
3. Specified rationale, goals and objectives
4. Objectives stated behaviorally
5. Objectives stated in terms of the learner
6. Objectives do include higher-order skills
7. Learners are informed of the objectives
8. Range and scope of content is adequate to achieve program's intents
9. Total program sequence
10. Preinstructional strategies:
    pretests
    advanced organizers
title at beginning of unit

11. Instructional text format
12. Concept learning
13. Vocabulary appropriate for learner
14. Graphics embedded in content
15. Demonstration of the exercise
17. Instructions clearly stated for the student
18. Evaluation components
19. Record-keeping device

Necessary for design of courseware:

1. Curriculum role:
   Adjunct
   Mainline
   Management only

2. Mode of interaction
   Drill and Practice
   Tutorial
   Game
   Simulation
   Problem Solving
   Exploration

3. Student sequence
   Nonlinear
   Varied by teacher/student
25

4. Instructional text format
5. Graphics embedded in content animation
6. Graphics used appropriately
7. Cues and/or prompts used
8. Action occurring on the screen
9. User control
10. Computer-managed instruction
11. Feedback
12. Records stored on magnetic devices for future retrieval
13. Program content designed to be altered
14. Random generation
15. Packaging designed for component parts
16. Teacher's manual and student manual
17. Technical design which allows:
   Quick response time
   Quick loading time

David Savitsky (1984) recommended a set of criteria which should be considered in the development of instructional software:

1. Each educational software program must have significant new content or skill-producing strategies.
2. Each educational software program must be motivating and offer some challenge.

3. Each educational software program must guarantee the learner an emotionally healthy and appropriate learning environment.

4. Each educational software must have educational objectives that are carefully chosen and clearly stated.

5. Each educational software program must be subjected to rigorous field testing with appropriate learners.

6. Each educational software program must provide the learner with ease of control over the elements of the learning experience.

7. Each educational software program must provide a trail to permit the teacher/adult to monitor and review the learner's activities.

8. Each educational software program must provide clear documentation about all aspects of the programs for the learner and the teacher.

9. Each educational software program must be accurate in the presentation of all facts, and precise in the use of spelling, grammar, and usage.
10. Each educational software program must be free of personal abuse, sarcasm, derogatory, or sexist/racist remarks.

More subjective criteria are provided by Hakansson (1981), where he suggested that the reviewer should "use your own judgment." Specific suggestions are made that evaluation should proceed to determine where activities are appropriate to the concepts to be taught, are appropriate for the age of the student user, call for reasonable tasks on the part of the user, serve as an appropriate medium for the content, and are attractive and interesting, ease of use, and presentation of a concept in a harmonious and well-balanced way.

Czechowicz (as cited in Steffin, 1983) suggested that a well-designed educational program should: (1) assume the user is naive; (2) include user/teacher documentation; (3) provide branching routines; (4) be userproofed; (5) recognize the need to escape or pause mid-exercise; (6) use good language; (7) give control of presentation rate to the user; (8) contain descriptive menus; (9) provide immediate non-judgmental responses; (10) reinforce correct responses; (11) be not merely page turning; (12) focus on defined objectives; (13) make appropriate use of graphics; and (14) contain screen displays designed for ease of viewing.
The purpose of instruction, in general, is to facilitate learning. Gagne and Briggs (1974) have identified "instructional events" which describe how instruction is accomplished. These events or components of instruction can provide a framework for classifying characteristics of instructional programs:

1. Gaining attention
2. Informing the learner of the objective
3. Stimulating recall of prerequisite learnings
4. Presenting the stimulus material
5. Providing "learning guidance"
6. Eliciting the performance
7. Providing feedback about performance correctness
8. Assessing the performance
9. Enhancing retention and transfer

He notes that not all of these events are necessarily found in every instructional program.

Wade (1980) identified five fundamental characteristics of good learning situations which can facilitate evaluating computer instructional programs:

1. The learning must be right
2. The learner must be ready
3. Learning needs to be managed or facilitated
4. Assimilation must be practicable
5. Learning must be efficient
Steinberg (1983) proposed three categories which are essential for a complete review of computer courseware: (1) suitability of the lesson for the intended population; (2) utility of implementation of unique features of CAI; and (3) observations of students users.

Coburn, Kelman, Roberts, Snyder, Watt, and Weiner (as cited in Muller, 1985) suggested that there be four broad areas of concern when evaluating a program: (1) program content—the suitability of materials for the students and the objectives, and the accuracy and significance of the content; (2) pedagogy—the nature of a program's feedback, the program developer's assumptions of learning, and the types of learning modes used; (3) program operation—the control that users have when using the program, the program's quality and the quality of the documentation; and (4) student outcome—the degree to which students learn what the program intends to teach, and the effectiveness of the program compared to non-computer-assisted instruction in the same area (Muller, 1985).

Owston proposed a criterion-based alternative to relatively subjective, comparative software evaluation schemes—the York Educational Software Evaluation Scales (YESSES). It examines educational software along four dimensions: pedagogical content, which refers to the
knowledge and skills the software purports to teach; instructional presentation, which addresses how well the software takes advantage of the unique capabilities of the microcomputer in presenting the content; documentation, which includes information and supporting materials on how to use the software both from a technical and a pedagogical point of view; and the technical adequacy, which refers to the design of the software with respect to user inputs, software outputs, and system errors (Owston, 1985). For each dimension, there is a four-point, criterion-based scale with the points representing "exemplary" software, "desirable" software, "minimally acceptable" software, and "deficient" software. Each point on the scale is defined by a set of descriptors that provide typical characteristics of software that would be rated at that level (Dudley-Marling & Owston, 1987).

Dudley-Marling and Owston (1987) asserted that the most important criterion for software evaluation is the response of individual teachers and students to the software. If the program supports the teacher's educational goals and assists student learning, it is a worthwhile program, regardless of other evaluation. Teachers can use formal evaluations of educational software to guide their initial screening of software, but their final judgments should depend upon their
own observations of their students as the students interact with the software.

Some authors have described ways to evaluate courseware without experimental investigation (e.g., Chambers and Sprecher; Cohen; Fetter; Futrell and Geisert; Jay; Steffin; Steinberg, as cited in Criswell & Swezey, 1984). The Futrell and Geisert (1984) courseware evaluation is specially comprehensive and includes checklist items in five areas: (1) courseware purpose, goals, and objectives; (2) evaluation of testing; (3) lesson presentation items including instructional processes and, material presentation considerations; (4) lesson content items; and (5) computer course management items. The others assert that courseware evaluation should receive increased emphasis because the training ability of a device rests fundamentally on courseware quality. It is commonly accepted that effective training should be designed around scientific principles of learning. Thus, a courseware evaluation should at a minimum assess the degree to which those principles are reflected in courseware sequences.

A major failing in courseware development is the lack of attention paid to formative evaluation. Roblyer (1981) issued a warning to companies that neglected this important phase in courseware development. She proposed three major
categories of criteria for courseware development: essential characteristics, aesthetic characteristics, and differential characteristics. She argues that these criteria and related characteristics are fundamental to all good instruction, regardless of medium or intended use.

Jay (1983) focused on five human information processing abilities which cognitive psychologists would anticipate must be accounted for in order to develop good courseware. Criteria for the formative evaluation of courseware advanced by Jay include: memory and attention demands, language or text characteristics, graphics and visual processing, a cognitive model of a user, and feedback (Duquette, 1985).

Marshall (1983) indicated that current instructional materials evaluation models do not adequately evaluate microcomputer courseware/software. He proposed a comprehensive evaluation model which was developed by a panel of experts for use in evaluating educational courseware/software. The model consisted of 41 evaluation criteria classified according to whether they were courseware/software identifying criteria, content criteria, instructional criteria, technical criteria, or documentation criteria. Those criteria were listed below:

Identifying criteria:
1. What type(s) of peripherals are needed to run the courseware/software?
2. What type of program is the courseware/software?
3. What ability level/grade is listed for the courseware/software?
4. What is the program title?
5. What microcomputer brand or model will the program run on?
6. What is the program's version number?
7. How much memory is required?
8. How much does the software cost?
9. Is any special hardware required?
10. Is any special software required to run the courseware/software?

Content criteria:
1. Are the concepts and facts accurate?
2. Is the scope of the content appropriate?
3. Is the content sequenced properly?
4. Is the content free of race, sex, and other stereotypes?
5. Is content grade level appropriate?

Technical criteria:
1. Does the input uses common symbols?
2. Are the punctuation and grammar in the programs correct?
3. Is the program free of technical errors?
4. Is some form of indication given to show where input of student responses go?
5. Is the courseware/software user friendly?
6. Is the screen format clear and easy to read?
7. Is there a high quality of student/computer interaction?

Instructional criteria:
1. Does the software achieve its stated objective?
2. Is the content format at the appropriate level of difficulty?
3. Is feedback appropriate?
4. Does the program give correct answers or offer assistance at the appropriate time?
5. Are the responses to errors non-judgmental?
6. Are good motivational techniques used?
7. Is the branching appropriate to meet the user's needs?
8. Does the program capture the student's interest?
9. Does the program engage problem solving skills?
10. Is the student in control of the program?

Documentation criteria:
1. Are the teaching instructions clear and complete?
2. Are the student instructions clear and complete?
3. Are directions clear and complete?
4. Are the prerequisite skills necessary to run the program defined?
5. Is the overall documentation acceptable?
6. Are the objectives well defined?
7. Does the menu allow access to the software components?
8. Is stop/start/re-entry information available?

MicroSIFT is a federally funded national clearinghouse for microcomputer-based courseware evaluations and related information. The MicroSIFT guide was developed by the Northwest Regional Educational Laboratory's Computer Technology program and is designed to facilitate a thorough, in-depth courseware evaluation. It is well and highly structured with clear explanations of the evaluation criteria used. The guide has been field-tested nationwide, and this final version is both a useful tool for the reviewer and a model for identifying excellence in courseware (Lathrop & Goodson, 1983). The model identifies three categories of evaluation criteria which were listed below:

Content quality:
1. The content is accurate.
2. The content has educational value.
3. The content is free of race, ethnic, sex, and other stereotypes.

Instructional quality:
1. The purpose of the package is well-defined.
2. The package achieves its defined purpose.
3. Presentation of content is clear and logical.
4. The level of difficulty is appropriate for the target audience.
5. Graphics/color/sound are used for appropriate instructional reasons.
6. Use of the package is motivational.
7. The package effectively stimulates student creativity.
8. Feedback on student responses is effectively employed.
9. The learner controls the rate and sequence of presentation and review.
10. Instruction is integrated with previous student experience.
11. Learning is generalizable to an appropriate range of situations.

Technical quality:
1. The user support materials are comprehensive.
2. The user support materials are effective.
3. Information displays are effective.
4. Intended users can easily and independently operate the program.
5. Teachers can easily employ the package.
6. The program appropriately uses relevant computer capabilities.
7. The program is reliable in normal use.

The California Library Media Consortium for Classroom Evaluation of Microcomputer Courseware came into being in December, 1981. It consisted of fifty-three library specialists representing 23 California counties. The Consortium developed a set of criteria for the evaluation of educational courseware/software. Along with the criteria, the Consortium also requires that all evaluations must be based on actual classroom use. The model identifies three aspects of criteria: general design, ease of use, and content. Those criteria are listed below:

General design:
1. Creative, innovative, effective use of computer.
2. Well-organized curriculum design.
3. Free of programming errors.
4. Free of excessive competition or violence.
5. Free of racial, ethnic, or sex stereotypes.

Ease of use:
1. Simple and complete instructions.
2. Screens neat and attractive.
3. Speed and sequence of paging can be controlled.
4. Technically easy to operate.
5. Any sound is appropriate and can be turned off.

Content:
1. Factual material, grammar, and spelling are correct.
2. Word lists, problems, and speed can be modified.
3. Interest level, difficulty, typing, and vocabulary are appropriate.
4. Provides easier or harder material in response to performance.
5. Response to student success is positive, enjoyable, and appropriate.
6. Avoid clear graphics that make it "fun to fail".

Heck, Johnson and Kansky (1981) developed a courseware/software evaluation model with supportive guidelines for evaluating computerized instructional materials. Their evaluation criteria are divided into nine classifications:

1. Instructional range.
2. Instructional grouping for program use.
3. Execution time.
4. Program user(s).
5. User orientation: instructor's point of view.
6. User orientation: student's point of view.
7. Content motivation and instructional style.
8. Social characteristics.

The Educational Products Information Exchange (EPIE) institute and the Microcomputer Resource Center (MRC) combined efforts in an attempt to develop an effective educational courseware/software evaluation instrument. Such a need came with the recognition that "the development and implementation of software packages had not been commensurate with the scales of educational hardware" (p. 4). The evaluation format is narrative. The criteria used in the courseware/software evaluation instrument include the major components of intents, contents, methodology, means of evaluation, instructional design congruence, use considerations, summary, and recommendation (Marshall, 1983).

Lathrop and Goodson (1983) noted that in evaluating a program, the first question must still be "Does it run on my computer?" The next question usually should be, "Does this program use our equipment to meet specific curriculum needs and objectives in a creative manner that, represents a good investment of instructional funds and computer time?" This important issue is frequently ignored in evaluating courseware. She also lists a lengthy list of evaluation criteria not contained in any one evaluation form. Those
criteria include creativity, content, screen formatting, instructions, student response, program response to student, motivational devices, technical quality, documentation, and teacher ability.

Day (1984) considers courseware selection to be a very difficult and complex process. It is difficult for several reasons. First, the needs of the local situation greatly influence and often determine the criteria used in an evaluation. A criterion that is quite appropriate in one situation might not apply at all in another. Second, it is very difficult for evaluators to agree how to weigh (value) criteria. The weighing of criteria often takes place in the minds of the evaluators, resulting in evaluations that are ultimately subjective in nature. Third, since evaluation implies decision-making, there are always some competitive alternatives of courseware which should be considered in the decision process, but sometimes due to budgetary and time constraints, the selection process can not be made thoroughly.

Weighting Criteria

Just as it is necessary to use criteria in the evaluation process, so it is necessary to assign value to, or weigh, those criteria (Day, 1984). Page (1979) pointed
out that the problem of value is central to all serious decision-making. Even the much vaunted behavioral objectives are virtually worthless without some attached weighting to establish priority and resource allocation.

Needs assignment is closely linked to how one establishes and values criteria. That is, the determined needs can actually become or can strongly influence how one weighs criteria (Day, 1984).

Bitter and Camuse (1984) asserted that each criterion should be given a "weight" according to its relative importance. The weight factor can range from 0 to 5, though you can select to use a greater range of values. If selecting a weight factor of 0 for a certain criterion statement, you indicate that the criterion has "no" importance for your purposes. A weight factor of 1 indicates that the criterion has only a small amount of value, and a weight factor of 5 would show that the criterion has a large amount of value for your purposes.

Some other courseware evaluation services or models try to weigh the criteria used in their evaluations. The MicroSIFT evaluation with simple H (High) and L (Low). The evaluation form used by School Microware Reviews weighs each criterion with scales from 1 (lowest) to 5 (highest) by a team of experts. After assigning raw scores to the
criteria, an evaluator then uses a mathematical formula to determine a final, weighted numerical score. McGraw's Computer program Appraisal (CPA) model and Wholeben's (as cited in Day, 1984) MICROPIK model both incorporate local needs and values. For example, CPA uses a weighted needs analysis (Scriven, as cited in Day, 1984) in which evaluators weigh the important needs in their local situation. The final rating then reflects these weighted needs (Day, 1984).

The courseware evaluation services and models which incorporate weighted needs and/or weighted criteria into the evaluation process are likely to rate courseware in order to determine which is the "best" among alternatives.

Summary

Researches have shown that computer-assisted instruction (CAI) is an effective medium for improving academic skills in significantly less time than conventional classroom methods (Kulik, Bangert, & Williams, as cited in Perez & White, 1985). However, these earlier studies were concerned with larger and older computer systems rather than with microcomputer systems that are so widely available in classroom today (Jolicoeur & Berger, 1986).
Because CAI is a new technology of instruction, especially using microcomputers as an instructional tool, there is no unified set of criteria that definitely identifies "high quality" courseware/software. Many different kinds of evaluation checklists and instruments were developed by many different institutions. None of which contains the same set of criteria.

Most of the evaluation devices are subjective in nature. Subsequently, subjectivity can only serve to lower reliability among separate ratings. For example, Jolicoeur and Berger (1986) proposed meta-analysis method trying to identify studies which can meet three conditions: (1) the study must have measured the effects of an individual software program; (2) performance must have been measured by an objective test; (3) there must have been a control group. Only two of the 47 studies (Davis; Watkins & Abram, as cited in Jolicoeur & Berger, 1986) met all requirements for the proposed meta-analysis. In addition, they also found that there is a striking lack of agreement between EPIE and MicroSIFT concerning the quality of specific educational software (Jolicoeur & Berger, 1986).

Almost every evaluation checklist or instrument uses the same set of criteria to evaluate different types of courseware/software, e.g., drill/practice, tutorial, problem
solving, game, etc. In addition, there is little systematic effort to design evaluations that take into account the weighing of criteria (Caldwell, 1983; Day, 1984).

Because of time and cost constraints, most of the evaluators are using internal criteria to evaluate courseware/software. Hence, many evaluations are subjective and not reliable. To improve this drawback, some researchers, like Jolicoeur and Berger (1986), Caldwell (1983), suggest to apply external criteria as possible to accurately predict the effectiveness of courseware/software, especially during the development stage of CAI programs. However, the trade-off is apparent; in-depth evaluations, such as Computer Program Appraisal (CPA) and MICROPIK, which do include external criteria in their evaluations, are more costly in terms of time and money. Perhaps it is not cost-effective to expend so many resources to evaluate the mass of new or existing courseware. Therefore, evaluators need to determine not only which evaluation methods are most effective in accurately predicting the effectiveness of courseware, but they also need to determine the cost (time and money) of performing the evaluations (Day, 1984).
CHAPTER III. METHODOLOGY

This chapter contains a summary of the procedures adopted for this study. These procedures were presented in the following aspects:

1. Population and samples.
2. The development of the instrument.
3. Data collection.
4. Data analysis.

Population and Samples

According to Curriculum Information Center's School Directory of School Year 1983-1984, there are 832 elementary schools and secondary schools out of total 1,506 schools in the state of Iowa which were using microcomputers in classroom teaching. Those schools were identified as the population of this study. A stratified random sampling technique was employed to select the sample size of 301 schools from the population. The sample consisted of 166 elementary schools and 135 secondary schools.

Instrument Development

The literature review revealed the existense of many microcomputer software evaluation instruments and checklists which were developed by different institutions and
individuals. No unified evaluation criteria set can be found. Therefore, the evaluation criteria instrument used in this study attempts to accommodate those different sets of evaluation criteria. The instrument consists of three categories of evaluation criteria: content quality, instructional quality and technical quality. These are general divisions. The content quality contains eight evaluation criteria. The instructional quality contains 19 evaluation criteria. The technical quality contains 10 evaluation criteria. The instrument was developed by the researcher. For pre-pilot testing evaluation and revision the instrument was then submitted to 10 graduate students in the Department of Industrial Education and Technology and five committee members. After the first revision was made, the instrument (see Appendices A and B) was mailed to 11 elementary schools and nine secondary schools which were randomly selected from the population and not included in the sample of this study. Nine schools replied. In the mailing instrument, an open-ended format was applied to ask if any criterion need to be added to the instrument, or if any criterion in the instrument need to be cancelled, or if the criteria content should be changed. The second revision was made based on these responses. At the proposal meeting of this study, the second-revised instrument was submitted
to the committee members for the final discussion and revision. The final instrument is included in Appendix D.

Data Collection

On May 8, 1987, 301 investigation questionnaires (final instrument) (see Appendices C and D) were mailed to the principals of 301 schools which comprised the sample. The researcher asked the principal of each school to forward the questionnaire to the teacher who was the most experienced in using microcomputer softwares of drill/practice and educational game. On May 20, 1987, 89 questionnaires were returned. On May 22, 1987, 202 follow-up letters (see Appendix E) were mailed. On June 9, 1987, the follow-up deadline, 27 questionnaires were mailed back. On June 15, another three came in. The final response was 119 questionnaires. One school responding indicated that no microcomputer was available in its school. Two schools responded that no qualified teacher was available to answer the questionnaire. The final return rate was 40%.

Data Analysis

The Statistical Analysis System (SAS) (SAS Institute Inc., 1985) package was used to analyze the collected data of this study.
The data were presented and analyzed in two aspects:

**Descriptive analysis**

A descriptive analysis was used to determine the mean and standard deviation of the scale value for each evaluation criterion of the drill/practice and educational game software programs.

**Inferential analysis**

A split-plot analysis of variance (2 x 2) (Kirk, 1982) was applied to determine if a statistically significant difference existed between the scale values of evaluation criteria from elementary school teachers and secondary school teachers.

A split-plot analysis of variance (2 x 2) was applied to determine if a statistically significant difference existed between the scale values of evaluation criteria from drill/practice and educational game software programs.

A split-plot analysis of variance (2 x 2) was applied to determine if a significant interaction existed between the scale values of school teachers and software programs.

In addition, the scale values in each evaluation criteria category (content quality, instructional quality, and technical quality) were summed independently to test each hypothesis of this study.
CHAPTER IV. RESULTS AND FINDINGS

In this chapter, the survey results and findings of this study are presented. There are three sections of this chapter:

1. The results of survey response.
2. The answers to research questions.
3. The testing of research hypothesis.

Survey Response

A careful effort was made to increase the return rate of the questionnaire in the process of conducting the survey. To encourage sampled subjects to answer the questionnaire, a form requesting "COPY OF RESULTS REQUESTED" was included with the questionnaire. The first-time respondents numbered 89. The return rate was 29.57%. To increase the return rate, the first follow-up was made. To encourage the non-respondents to answer the questionnaire, a gift certificate of McDonald and Bonanza was mailed to the non-respondent with the chance of winning being 1:10. New questionnaires were mailed again to those who requested it. Thirty-three questionnaires were returned, which increased the return rate to 40.53%. The deadline of first follow-up was on June 1, 1987, which was very close to the end of the school semester. Therefore, a second follow-up was not conducted.
The total number of returned questionnaires was 122 and the return rate was 40.53% (see Table 1). One respondent indicated that there was no microcomputer available in his school. Two respondents answered that no qualified teacher could be found to answer the questionnaire. An additional five respondents reported that they had no experience in using drill/practice and/or educational game software programs. Those eight responses were excluded from the final data analysis.

Pilot Study of the Data Collection Instrument

A pilot study was conducted on the survey instrument (questionnaire) to identify problems which sampled subjects may have in completing the questionnaire. Twenty schools (eleven elementary schools and nine secondary schools) were randomly selected from the population excluding the sampled subjects. The questionnaire was mailed along with a self-addressed, stamped envelope. In addition to responding to the survey, school teachers were asked to complete information (questionnaire PART C) about the survey. The additional information collected included directions to check for clarity and time required to complete the survey, items that may have been confusing, or other comments about the questionnaire.
Nine schools (45%) returned pilot questionnaires. According to their responses, questionnaire directions were clear enough to complete the questionnaire. The average time to complete the survey was 14.1 minutes. Four responded that items 4 (p. 2), 15 (p. 4), 2 (p. 4), 4 (p. 4), and 7 (p. 5) (see Appendix B) were confusing. One teacher suggested to add one evaluation criterion to technical quality: must have record keeping built into program. The graduate committee of the researcher also suggested some changes of these items: 8 (p. 1), 4 (p. 2), 5 (p. 3), 9 (p. 3), 12 (p. 4), 15 (p. 4), 2 (p. 4), and 4 (p. 4). They also suggested adding two evaluation criteria to technical quality: The system provides a summary of student progress; Student responses are automatically recorded for teacher use. The final version of the survey survey instrument may be found in Appendix D.

Response of the Basic Data

The sex distribution of respondents is presented in Table 2. There were 49 males (43.36%), and 64 females (56.64%).

Table 3 presented the group of respondents by teaching level. The distribution was 59 (51.75%) for elementary school teachers, and 55 (48.25%) for secondary school teachers.
TABLE 1. The number and percent of respondents in sample

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respondent</td>
<td>122</td>
<td>40.53</td>
</tr>
<tr>
<td>Non-respondent</td>
<td>179</td>
<td>59.47</td>
</tr>
<tr>
<td>Total</td>
<td>301</td>
<td>100.00</td>
</tr>
</tbody>
</table>

TABLE 2. The sex of respondents

<table>
<thead>
<tr>
<th>Sex</th>
<th>N</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>49</td>
<td>43.36</td>
</tr>
<tr>
<td>Female</td>
<td>64</td>
<td>56.64</td>
</tr>
</tbody>
</table>

TABLE 3. The number and percent of school teachers classified by teaching level

<table>
<thead>
<tr>
<th>Level</th>
<th>N</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elementary school teachers</td>
<td>59</td>
<td>51.75</td>
</tr>
<tr>
<td>Secondary school teachers</td>
<td>55</td>
<td>48.25</td>
</tr>
</tbody>
</table>

The types of microcomputer software programs used in classroom teaching are presented in Table 4. The most
TABLE 4. Types of microcomputer software programs used in classroom teaching

<table>
<thead>
<tr>
<th>Type</th>
<th>Schools</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Drill/practice</td>
<td>96</td>
<td>17.08</td>
</tr>
<tr>
<td>2. Educational game</td>
<td>83</td>
<td>14.77</td>
</tr>
<tr>
<td>3. Word processing</td>
<td>80</td>
<td>14.23</td>
</tr>
<tr>
<td>4. Tutorial</td>
<td>65</td>
<td>11.57</td>
</tr>
<tr>
<td>5. Problem solving</td>
<td>63</td>
<td>11.21</td>
</tr>
<tr>
<td>6. Simulation</td>
<td>48</td>
<td>8.54</td>
</tr>
<tr>
<td>7. Utility</td>
<td>44</td>
<td>7.83</td>
</tr>
<tr>
<td>8. Management/</td>
<td>41</td>
<td>7.30</td>
</tr>
<tr>
<td>Administration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Information</td>
<td>20</td>
<td>3.56</td>
</tr>
<tr>
<td>10. Testing</td>
<td>13</td>
<td>2.31</td>
</tr>
<tr>
<td>11. Other</td>
<td>9</td>
<td>1.60</td>
</tr>
</tbody>
</table>

commonly used included drill/practice, educational game, word processing, tutorial, and problem solving. The three most popular types of software programs were drill/practice, educational game and word processing.

Table 5 revealed the number of years of teaching experience of school teachers. Most of the teachers (88.60%) had more than five years of teaching experience. Approximately 25% had more than 20 years of teaching experience.
TABLE 5. The number of years of teaching experience

<table>
<thead>
<tr>
<th>Years</th>
<th>Teachers</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5</td>
<td>13</td>
<td>11.40</td>
</tr>
<tr>
<td>6-10</td>
<td>25</td>
<td>21.93</td>
</tr>
<tr>
<td>11-15</td>
<td>26</td>
<td>22.81</td>
</tr>
<tr>
<td>16-20</td>
<td>21</td>
<td>18.42</td>
</tr>
<tr>
<td>More than 21</td>
<td>29</td>
<td>25.44</td>
</tr>
</tbody>
</table>

TABLE 6. The number of years experience with computers

<table>
<thead>
<tr>
<th>Years</th>
<th>Teachers</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5</td>
<td>63</td>
<td>55.75</td>
</tr>
<tr>
<td>6-10</td>
<td>43</td>
<td>38.05</td>
</tr>
<tr>
<td>11-15</td>
<td>1</td>
<td>0.89</td>
</tr>
<tr>
<td>16-20</td>
<td>5</td>
<td>4.42</td>
</tr>
<tr>
<td>More than 20</td>
<td>1</td>
<td>0.89</td>
</tr>
</tbody>
</table>

The number of years experience with computers was presented in Table 6. About half of the teachers had more than five years of experience with computers.

Table 7 revealed that most teachers (93.81%) have had experience in using drill/practice software programs.
TABLE 7. The number and percent of teachers with experience in using drill/practice software programs

<table>
<thead>
<tr>
<th>Experience</th>
<th>Teachers</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>110</td>
<td>96.49</td>
</tr>
<tr>
<td>No</td>
<td>4</td>
<td>3.51</td>
</tr>
</tbody>
</table>

TABLE 8. The number and percent of teachers with experience in using educational software programs

<table>
<thead>
<tr>
<th>Experience</th>
<th>Teachers</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>106</td>
<td>93.81</td>
</tr>
<tr>
<td>No</td>
<td>7</td>
<td>6.19</td>
</tr>
</tbody>
</table>

As indicated in Table 8, most teachers (93.81%) had experience in using educational game software programs.

As presented in Table 9, most teachers (94.73%) thought that computer software is most useful when it is used as a supplement to other methods of instruction.

Results of the Hypothesis Testing

Hypothesis 1

There was no significant difference between the scale values of content quality evaluation criteria from responses of elementary school teachers and secondary school teachers.
TABLE 9. Situation in which computer software is most useful

<table>
<thead>
<tr>
<th>Situation</th>
<th>Response</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supplement to other methods of instruction</td>
<td>108</td>
<td>94.73</td>
</tr>
<tr>
<td>Means for developing computer skills</td>
<td>4</td>
<td>3.51</td>
</tr>
<tr>
<td>Substitute for other methods of instruction</td>
<td>1</td>
<td>0.88</td>
</tr>
<tr>
<td>Currently isn't very useful</td>
<td>1</td>
<td>0.88</td>
</tr>
</tbody>
</table>

The results of hypothesis 1 are presented in Table 10. The statistical model had an F-value of 8.76, a statistical significance of < .001, accounted for much of the dependent variable's variability. The data provided evidence that null hypothesis 1 can be rejected at $\alpha = 0.05$ level.

Hypothesis 2

There was no significant difference between the scale values of instructional quality evaluation criteria from responses of elementary school teachers and secondary school teachers.

Table 11 revealed that the statistical model with an F-value of 11.24, a statistical significance of < .001, accounted for most of the dependent variable's variability.
TABLE 10. The split-plot ANOVA of teacher groups and types of software programs for content quality evaluation criteria

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>F Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>111</td>
<td>13457.12</td>
<td>121.24</td>
<td>8.76</td>
<td>&lt; .001**</td>
</tr>
<tr>
<td>Error</td>
<td>101</td>
<td>1397.50</td>
<td>13.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected total</td>
<td>212</td>
<td>14854.63</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>F Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gr</td>
<td>1</td>
<td>883.10</td>
<td>883.10</td>
<td>8.02</td>
<td>&lt; .005</td>
</tr>
<tr>
<td>Teid (Gr)</td>
<td>108</td>
<td>11893.12</td>
<td>110.12</td>
<td>7.96</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>545.21</td>
<td>545.21</td>
<td>39.40</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>Gr*B</td>
<td>1</td>
<td>42.15</td>
<td>42.15</td>
<td>3.05</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>B*Teid (Gr)</td>
<td>101</td>
<td>1397.50</td>
<td>13.84</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mean(DP)</th>
<th>SD(DP)</th>
<th>Mean(EG)</th>
<th>SD(EG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ES</td>
<td>45.76</td>
<td>6.66</td>
<td>43.20</td>
</tr>
<tr>
<td>SS</td>
<td>42.43</td>
<td>8.03</td>
<td>38.16</td>
</tr>
</tbody>
</table>

Gr : ES - Elementary School Teachers.
SS - Secondary School Teachers.
B : DP - Drill/Practice Software Programs.
EG - Educational Game Software Programs.
Teid : Teachers within Groups.

**Significant at the p < .01 level.

Based on the results in Table 11, hypothesis 2 was retained at \( \alpha = 0.05 \) level.
TABLE 11. The split-plot ANOVA of teacher groups and types of software programs for instructional quality evaluation criteria

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>F Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>103</td>
<td>61447.09</td>
<td>596.57</td>
<td>11.24</td>
<td>&lt; .001**</td>
</tr>
<tr>
<td>Error</td>
<td>92</td>
<td>4881.83</td>
<td>53.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected total</td>
<td>195</td>
<td>66328.92</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>F Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gr</td>
<td>1</td>
<td>1585.76</td>
<td>1585.76</td>
<td>2.74</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>Teid (Gr)</td>
<td>100</td>
<td>57858.85</td>
<td>578.59</td>
<td>10.90</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>1516.79</td>
<td>1516.79</td>
<td>28.58</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>Gr*C</td>
<td>1</td>
<td>113.38</td>
<td>113.38</td>
<td>2.14</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>C*Teid (Gr)</td>
<td>92</td>
<td>4881.83</td>
<td>53.06</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>Mean(DP)</th>
<th>SD(DP)</th>
<th>Mean(EG)</th>
<th>SD(EG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ES</td>
<td>102.62</td>
<td>17.68</td>
<td>97.83</td>
<td>19.69</td>
</tr>
<tr>
<td>SS</td>
<td>98.36</td>
<td>14.89</td>
<td>90.68</td>
<td>19.83</td>
</tr>
</tbody>
</table>

Gr : ES - Elementary School Teachers.
SS - Secondary School Teachers.
C : DP - Drill/Practice Software Programs.
EG - Educational Game Software Programs.
Teid : Teachers within Groups.

**Significant at the p < .01 level.

Hypothesis 3
There was no significant difference between the scale values of technical quality evaluation criteria from responses of elementary school teachers and secondary school teachers.
TABLE 12. The split-plot ANOVA of teacher groups and types of software programs for technical quality evaluation criteria

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>F Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>113</td>
<td>18430.21</td>
<td>163.10</td>
<td>9.19</td>
<td>&lt; .001**</td>
</tr>
<tr>
<td>Error</td>
<td>105</td>
<td>1862.66</td>
<td>17.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected total</td>
<td>218</td>
<td>20292.88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gr</td>
<td>1</td>
<td>229.48</td>
<td>229.48</td>
<td>1.44</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>Teid (Gr)</td>
<td>110</td>
<td>17535.79</td>
<td>159.42</td>
<td>8.99</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>586.11</td>
<td>586.11</td>
<td>33.04</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>Gr*D</td>
<td>1</td>
<td>31.35</td>
<td>31.35</td>
<td>1.77</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>D*Teid (Gr)</td>
<td>105</td>
<td>1862.66</td>
<td>17.74</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Mean(DP)</th>
<th>SD(DP)</th>
<th>Mean(EG)</th>
<th>SD(EG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ES</td>
<td>56.55</td>
<td>9.59</td>
<td>54.02</td>
<td>10.01</td>
</tr>
<tr>
<td>SS</td>
<td>55.26</td>
<td>8.16</td>
<td>51.06</td>
<td>10.08</td>
</tr>
</tbody>
</table>

Gr : ES - Elementary School Teachers.
SS - Secondary School Teachers.
D : DP - Drill/Practice Software Programs.
EG - Educational Game Software Programs.
Teid : Teachers within Groups.

**Significant at the p < .01 level.

The results of hypothesis 3 are presented in Table 12. As a whole, The model accounted for most of the dependent variable's variability. An F-value of 1.44, with a statistical significance of > .05, was found by the two-way
analysis of variance for the difference between elementary and secondary school teachers. Therefore, based on the results in Table 12, hypothesis 3 was not rejected at $\alpha = 0.05$ level.

**Hypothesis 4**

There was no significant difference between the scale values of content quality evaluation criteria from drill/practice and educational game software programs.

Based on the findings of Table 10, hypothesis 4 was rejected at $\alpha = 0.05$ level.

**Hypothesis 5**

There was no significant difference between the scale values of instructional quality evaluation criteria from drill/practice and educational game software programs.

Table 11 provided evidence that hypothesis 5 may be rejected at $\alpha = 0.05$ level.

**Hypothesis 6**

There was no significant difference between the scale values of technical quality evaluation criteria from drill/practice and educational game software programs.

Based on the findings of Table 12, hypothesis 6 was rejected at $\alpha = 0.05$ level.
Hypothesis 7
There was no interaction between the scale values of content quality evaluation criteria from school teachers and software programs.

Table 10 provided evidence that hypothesis 7 may be retained at $\alpha = 0.05$ level.

Hypothesis 8
There was no interaction between the scale values of instructional quality evaluation criteria from school teachers and software programs.

Based on the findings of Table 11, hypothesis 8 should be retained at $\alpha = 0.05$ level.

Hypothesis 9
There was no interaction between the scale values of technical quality evaluation criteria from school teachers and software programs.

Table 12 provided evidence that hypothesis 9 should be retained at $\alpha = 0.05$ level.

Results of the Research Questions

Table 13 presents the item importance priority, mean, and standard deviation of evaluation criteria from combined-group for drill/practice software programs. With respect to
<table>
<thead>
<tr>
<th>Item importance priority</th>
<th>Evaluation criteria for drill/practice software programs</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The content is accurate</td>
<td>6.74</td>
<td>0.65</td>
</tr>
<tr>
<td>2</td>
<td>The program is reliable in normal use</td>
<td>6.49</td>
<td>0.83</td>
</tr>
<tr>
<td>3</td>
<td>The content has educational value</td>
<td>6.48</td>
<td>0.89</td>
</tr>
<tr>
<td>4</td>
<td>Presentation of content is clear, logical and well-organized</td>
<td>6.45</td>
<td>0.78</td>
</tr>
<tr>
<td>5</td>
<td>Teachers can easily employ the package</td>
<td>6.39</td>
<td>0.94</td>
</tr>
<tr>
<td>6</td>
<td>Intended users can independently operate the program</td>
<td>6.27</td>
<td>1.05</td>
</tr>
<tr>
<td>7</td>
<td>The level of difficulty is appropriate for the target audience</td>
<td>6.19</td>
<td>0.97</td>
</tr>
<tr>
<td>8</td>
<td>The program encourages learning the subject material</td>
<td>6.09</td>
<td>1.15</td>
</tr>
<tr>
<td>9</td>
<td>The package achieves its defined purpose</td>
<td>6.07</td>
<td>1.16</td>
</tr>
<tr>
<td>10</td>
<td>The content is free of race, ethnic, sex, and other stereotypes</td>
<td>6.06</td>
<td>1.63</td>
</tr>
<tr>
<td>11</td>
<td>High student involvement and interaction</td>
<td>6.04</td>
<td>1.23</td>
</tr>
<tr>
<td>12</td>
<td>Feedback on student responses is effectively employed</td>
<td>5.96</td>
<td>1.25</td>
</tr>
<tr>
<td>13</td>
<td>The purpose of the package is well defined</td>
<td>5.83</td>
<td>1.24</td>
</tr>
<tr>
<td>14</td>
<td>Amount of learning justifies time spent by users</td>
<td>5.79</td>
<td>1.56</td>
</tr>
</tbody>
</table>
TABLE 13 (Continued)

<table>
<thead>
<tr>
<th>Item importance priority</th>
<th>Evaluation criteria for drill/practice software programs</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>The user support materials are effective in assisting the user to run the program</td>
<td>5.77</td>
<td>1.49</td>
</tr>
<tr>
<td>16</td>
<td>Follows sound educational theory</td>
<td>5.74</td>
<td>1.42</td>
</tr>
<tr>
<td>17</td>
<td>Follows sound educational techniques</td>
<td>5.70</td>
<td>1.39</td>
</tr>
<tr>
<td>18</td>
<td>The content is complete</td>
<td>5.66</td>
<td>1.78</td>
</tr>
<tr>
<td>19</td>
<td>Use of the package is motivational</td>
<td>5.62</td>
<td>1.43</td>
</tr>
<tr>
<td>20</td>
<td>The user can control the rate of presentation and review</td>
<td>5.60</td>
<td>1.73</td>
</tr>
<tr>
<td>21</td>
<td>Information displays are effective</td>
<td>5.51</td>
<td>1.56</td>
</tr>
<tr>
<td>22</td>
<td>The system provides a summary of student progress</td>
<td>5.40</td>
<td>1.65</td>
</tr>
<tr>
<td>23</td>
<td>The content is free of excessive violence</td>
<td>5.38</td>
<td>2.07</td>
</tr>
<tr>
<td>24</td>
<td>The user support materials are comprehensive</td>
<td>5.30</td>
<td>1.76</td>
</tr>
<tr>
<td>25</td>
<td>The program uses a well-organized curriculum design</td>
<td>5.20</td>
<td>1.85</td>
</tr>
<tr>
<td>26</td>
<td>The program appropriately uses relevant computer capabilities</td>
<td>5.18</td>
<td>1.95</td>
</tr>
<tr>
<td>27</td>
<td>Instruction is integrated with previous student experience</td>
<td>4.96</td>
<td>1.95</td>
</tr>
<tr>
<td>28</td>
<td>Student responses are automatically recorded for teacher use</td>
<td>4.96</td>
<td>2.00</td>
</tr>
<tr>
<td>29</td>
<td>The user can control the sequence of presentation and review</td>
<td>4.75</td>
<td>1.99</td>
</tr>
<tr>
<td>30</td>
<td>The visual display is attractive, exciting, and absorbing</td>
<td>4.73</td>
<td>1.97</td>
</tr>
</tbody>
</table>
### TABLE 13 (Continued)

<table>
<thead>
<tr>
<th>Item importance priority</th>
<th>Evaluation criteria for drill/practice software programs</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>Graphics is used for appropriate instructional reasons</td>
<td>4.61</td>
<td>1.81</td>
</tr>
<tr>
<td>32</td>
<td>The content of the program is available for inspection and/or change</td>
<td>4.49</td>
<td>2.25</td>
</tr>
<tr>
<td>33</td>
<td>Learning is generalizable to an appropriate range of situations</td>
<td>4.26</td>
<td>1.99</td>
</tr>
<tr>
<td>34</td>
<td>The package effectively stimulates student creativity</td>
<td>3.97</td>
<td>2.43</td>
</tr>
<tr>
<td>35</td>
<td>Color is used for appropriate instructional reasons</td>
<td>3.75</td>
<td>2.17</td>
</tr>
<tr>
<td>36</td>
<td>The content is free of excessive competition</td>
<td>3.66</td>
<td>2.26</td>
</tr>
<tr>
<td>37</td>
<td>Sound is used for appropriate instructional reasons</td>
<td>3.35</td>
<td>2.22</td>
</tr>
</tbody>
</table>

The means, 11 evaluation criteria were rated above 6.00, 15 were rated from 5.00 to 5.99, seven were rated from 4.00 to 4.99, and four were rated below 4.00.

The item importance priority, mean, and standard deviation of evaluation criteria for the combined-group of educational game software programs are presented in Table 14.

The eight evaluation criteria were rated above 6.00, 15 were rated from 5.00 to 5.99, seven were rated from 4.00 to 4.99, and seven were rated below 4.00.
TABLE 14. The item importance priority, mean, and standard deviation of evaluation criteria from combined-group for educational game software programs

<table>
<thead>
<tr>
<th>Item importance priority</th>
<th>Evaluation criteria for educational game software programs</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The content is accurate</td>
<td>6.40</td>
<td>1.29</td>
</tr>
<tr>
<td>2</td>
<td>The program is reliable in normal use</td>
<td>6.40</td>
<td>0.94</td>
</tr>
<tr>
<td>3</td>
<td>Teachers can easily employ the package</td>
<td>6.22</td>
<td>1.32</td>
</tr>
<tr>
<td>4</td>
<td>Intended users can independently operate the program</td>
<td>6.15</td>
<td>1.11</td>
</tr>
<tr>
<td>5</td>
<td>Presentation of content is clear, logical, and well-organized</td>
<td>6.15</td>
<td>0.97</td>
</tr>
<tr>
<td>6</td>
<td>High student involvement and interaction</td>
<td>6.08</td>
<td>1.20</td>
</tr>
<tr>
<td>7</td>
<td>The content is free of race, ethnic, sex, and other stereotypes</td>
<td>6.03</td>
<td>1.75</td>
</tr>
<tr>
<td>8</td>
<td>The content has educational value</td>
<td>6.02</td>
<td>1.46</td>
</tr>
<tr>
<td>9</td>
<td>The level of difficulty is appropriate for the target audience</td>
<td>5.94</td>
<td>1.24</td>
</tr>
<tr>
<td>10</td>
<td>The package achieves its defined purpose</td>
<td>5.72</td>
<td>1.46</td>
</tr>
<tr>
<td>11</td>
<td>The user support materials are effective in assisting the user to run the program</td>
<td>5.59</td>
<td>1.68</td>
</tr>
<tr>
<td>12</td>
<td>Use of the package is motivational</td>
<td>5.45</td>
<td>1.48</td>
</tr>
<tr>
<td>13</td>
<td>The purpose of the package is well defined</td>
<td>5.37</td>
<td>1.61</td>
</tr>
<tr>
<td>14</td>
<td>Information displays are effective</td>
<td>5.36</td>
<td>1.78</td>
</tr>
<tr>
<td>Item</td>
<td>Evaluation criteria for educational game software programs</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>------</td>
<td>----------------------------------------------------------</td>
<td>------</td>
<td>----</td>
</tr>
<tr>
<td>15</td>
<td>The content is free of excessive violence</td>
<td>5.34</td>
<td>2.08</td>
</tr>
<tr>
<td>16</td>
<td>Follows sound educational theory</td>
<td>5.29</td>
<td>1.96</td>
</tr>
<tr>
<td>17</td>
<td>Feedback on student responses is effectively employed</td>
<td>5.28</td>
<td>1.82</td>
</tr>
<tr>
<td>18</td>
<td>The content is complete</td>
<td>5.17</td>
<td>2.16</td>
</tr>
<tr>
<td>19</td>
<td>The program encourages learning the subject material</td>
<td>5.12</td>
<td>2.07</td>
</tr>
<tr>
<td>20</td>
<td>The program appropriately use relevant computer capabilities</td>
<td>5.11</td>
<td>1.98</td>
</tr>
<tr>
<td>21</td>
<td>Follows sound educational techniques</td>
<td>5.07</td>
<td>1.96</td>
</tr>
<tr>
<td>22</td>
<td>Amount of learning justifies time spent by users</td>
<td>5.01</td>
<td>2.11</td>
</tr>
<tr>
<td>23</td>
<td>The visual display is attractive, exciting, and absorbing</td>
<td>5.00</td>
<td>1.77</td>
</tr>
<tr>
<td>24</td>
<td>The user support materials are comprehensive</td>
<td>4.88</td>
<td>2.00</td>
</tr>
<tr>
<td>25</td>
<td>The user can control the rate of presentation and review</td>
<td>4.86</td>
<td>2.07</td>
</tr>
<tr>
<td>26</td>
<td>The package effectively stimulates student creativity</td>
<td>4.83</td>
<td>2.11</td>
</tr>
<tr>
<td>27</td>
<td>Graphics is used for appropriate instructional reasons</td>
<td>4.63</td>
<td>1.85</td>
</tr>
<tr>
<td>28</td>
<td>The program uses a well-organized curriculum design</td>
<td>4.43</td>
<td>2.20</td>
</tr>
<tr>
<td>29</td>
<td>The system provides a summary of student progress</td>
<td>4.30</td>
<td>2.15</td>
</tr>
</tbody>
</table>
Table 14 (Continued)

<table>
<thead>
<tr>
<th>Item importance priority</th>
<th>Evaluation criteria for educational game software programs</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The user can control the sequence of presentation and review</td>
<td>4.05</td>
<td>2.16</td>
</tr>
<tr>
<td>31</td>
<td>Instruction is integrated with previous student experience</td>
<td>3.98</td>
<td>2.20</td>
</tr>
<tr>
<td>32</td>
<td>Learning is generalizable to an appropriate instructional reasons</td>
<td>3.93</td>
<td>2.18</td>
</tr>
<tr>
<td>33</td>
<td>Color is used for appropriate instructional reasons</td>
<td>3.73</td>
<td>2.27</td>
</tr>
<tr>
<td>34</td>
<td>Student responses are automatically recorded for teacher use</td>
<td>3.65</td>
<td>2.47</td>
</tr>
<tr>
<td>35</td>
<td>The content of the program is available for inspection and/or change</td>
<td>3.60</td>
<td>2.51</td>
</tr>
<tr>
<td>36</td>
<td>Sound is used for appropriate instructional reasons</td>
<td>3.31</td>
<td>2.21</td>
</tr>
<tr>
<td>37</td>
<td>The content is free of excessive competition</td>
<td>3.15</td>
<td>2.31</td>
</tr>
</tbody>
</table>

Table 15 presents the item importance priority, mean, and standard deviation of evaluation criteria from elementary school teachers for drill/practice software programs.

Table 15 also indicated that 12 evaluation criteria were rated above 6.00, 18 were rated from 5.00 to 5.99, six were rated from 4.00 to 4.99, and only one was rated below 4.00.
TABLE 15. The item importance priority, mean, and standard deviation of evaluation criteria from elementary school teachers for drill/practice software programs

<table>
<thead>
<tr>
<th>Item importance priority</th>
<th>Evaluation criteria for drill/practice software programs</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The content is accurate</td>
<td>6.73</td>
<td>0.71</td>
</tr>
<tr>
<td>2</td>
<td>Presentation of content is clear, logical, and well-organized</td>
<td>6.52</td>
<td>0.75</td>
</tr>
<tr>
<td>3</td>
<td>The content has educational value</td>
<td>6.49</td>
<td>0.91</td>
</tr>
<tr>
<td>4</td>
<td>Teachers can easily employ the package</td>
<td>6.46</td>
<td>1.08</td>
</tr>
<tr>
<td>5</td>
<td>Intended users can independently operate the program</td>
<td>6.42</td>
<td>0.85</td>
</tr>
<tr>
<td>6</td>
<td>The program is reliable in normal use</td>
<td>6.31</td>
<td>1.03</td>
</tr>
<tr>
<td>7</td>
<td>The level of difficulty is appropriate for the target audience</td>
<td>6.31</td>
<td>1.03</td>
</tr>
<tr>
<td>8</td>
<td>The program encourages learning the subject materials</td>
<td>6.20</td>
<td>1.03</td>
</tr>
<tr>
<td>9</td>
<td>High student involvement and interaction</td>
<td>6.19</td>
<td>1.13</td>
</tr>
<tr>
<td>10</td>
<td>The package achieves its defined purpose</td>
<td>6.05</td>
<td>1.29</td>
</tr>
<tr>
<td>11</td>
<td>The content is free of race, ethnic, sex, and other stereotypes</td>
<td>6.03</td>
<td>1.58</td>
</tr>
<tr>
<td>12</td>
<td>Feedback on student responses is effectively employed</td>
<td>6.02</td>
<td>1.30</td>
</tr>
<tr>
<td>13</td>
<td>The content is free of excessive violence</td>
<td>5.95</td>
<td>1.43</td>
</tr>
<tr>
<td>14</td>
<td>Follows sound educational techniques</td>
<td>5.88</td>
<td>1.38</td>
</tr>
<tr>
<td>Item priority</td>
<td>Evaluation criteria for drill/mean importance</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>15</td>
<td>Amount of learning justifies time spent by users</td>
<td>5.84</td>
<td>1.54</td>
</tr>
<tr>
<td>16</td>
<td>The purpose of the package is well defined</td>
<td>5.83</td>
<td>1.39</td>
</tr>
<tr>
<td>17</td>
<td>Follows sound educational theory</td>
<td>5.81</td>
<td>1.59</td>
</tr>
<tr>
<td>18</td>
<td>Use of the package is motivational</td>
<td>5.80</td>
<td>1.42</td>
</tr>
<tr>
<td>19</td>
<td>The user support materials are effective in assisting the user to run the program</td>
<td>5.76</td>
<td>1.54</td>
</tr>
<tr>
<td>20</td>
<td>The content is complete</td>
<td>5.69</td>
<td>1.79</td>
</tr>
<tr>
<td>21</td>
<td>The user can control the rate of presentation and review</td>
<td>5.47</td>
<td>1.78</td>
</tr>
<tr>
<td>22</td>
<td>The system provides a summary of student progress</td>
<td>5.47</td>
<td>1.72</td>
</tr>
<tr>
<td>23</td>
<td>Information displays are effective</td>
<td>5.42</td>
<td>1.88</td>
</tr>
<tr>
<td>24</td>
<td>The program uses a well-organized curriculum design</td>
<td>5.24</td>
<td>1.91</td>
</tr>
<tr>
<td>25</td>
<td>Student responses are automatically recorded for teacher use</td>
<td>5.24</td>
<td>1.81</td>
</tr>
<tr>
<td>26</td>
<td>The user support materials are comprehensive</td>
<td>5.20</td>
<td>1.85</td>
</tr>
<tr>
<td>27</td>
<td>The program appropriately uses relevant computer capabilities</td>
<td>5.15</td>
<td>2.10</td>
</tr>
<tr>
<td>28</td>
<td>The visual display is attractive, exciting, and absorbing</td>
<td>5.07</td>
<td>1.76</td>
</tr>
<tr>
<td>29</td>
<td>The content of the program is available for inspection and/or change</td>
<td>5.02</td>
<td>1.90</td>
</tr>
</tbody>
</table>
TABLE 15 (Continued)

<table>
<thead>
<tr>
<th>Item importance priority</th>
<th>Evaluation criteria for drill/</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Instruction is integrated with previous student experience</td>
<td>5.02</td>
<td>2.00</td>
</tr>
<tr>
<td>30</td>
<td>Graphics is used for appropriate instructional reasons</td>
<td>4.81</td>
<td>1.87</td>
</tr>
<tr>
<td>31</td>
<td>The user can control the sequence of presentation and review</td>
<td>4.71</td>
<td>1.99</td>
</tr>
<tr>
<td>32</td>
<td>Learning is generalizable to an appropriate range of situations</td>
<td>4.55</td>
<td>1.87</td>
</tr>
<tr>
<td>33</td>
<td>The package effectively stimulates student creativity</td>
<td>4.50</td>
<td>2.17</td>
</tr>
<tr>
<td>34</td>
<td>Color is used for appropriate instructional reasons</td>
<td>4.34</td>
<td>2.12</td>
</tr>
<tr>
<td>35</td>
<td>The content is free of excessive competition</td>
<td>4.05</td>
<td>2.09</td>
</tr>
<tr>
<td>36</td>
<td>Sound is used for appropriate instructional reasons</td>
<td>3.62</td>
<td>2.28</td>
</tr>
</tbody>
</table>

Table 16 presents the item importance priority, mean, and standard deviation of evaluation criteria from elementary school teachers for educational game software programs.

Table 16 also revealed that eight evaluation criteria were rated above 6.00, 16 were rated from 5.00 to 5.99, 11 were rated from 4.00 to 4.99, and two were rated below 4.00.
TABLE 16. The item importance priority, mean, and standard deviation of evaluation criteria from elementary school teachers for educational game software programs

<table>
<thead>
<tr>
<th>Item importance priority</th>
<th>Evaluation criteria for educational game software programs</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The content is accurate</td>
<td>6.41</td>
<td>1.23</td>
</tr>
<tr>
<td>2</td>
<td>Teachers can easily employ the package</td>
<td>6.39</td>
<td>1.20</td>
</tr>
<tr>
<td>3</td>
<td>The program is reliable in normal use</td>
<td>6.33</td>
<td>1.04</td>
</tr>
<tr>
<td>4</td>
<td>Intended users can independently operate the program</td>
<td>6.32</td>
<td>0.96</td>
</tr>
<tr>
<td>5</td>
<td>Presentation of content is clear, logical, and well-organized</td>
<td>6.27</td>
<td>1.08</td>
</tr>
<tr>
<td>6</td>
<td>The content has educational value</td>
<td>6.21</td>
<td>1.35</td>
</tr>
<tr>
<td>7</td>
<td>High student involvement and interaction</td>
<td>6.19</td>
<td>1.02</td>
</tr>
<tr>
<td>8</td>
<td>The content is free of race, ethnic, sex, and other stereotypes</td>
<td>6.05</td>
<td>1.64</td>
</tr>
<tr>
<td>9</td>
<td>The level of difficulty is appropriate for the target audience</td>
<td>5.98</td>
<td>1.41</td>
</tr>
<tr>
<td>10</td>
<td>The content is free of excessive violence</td>
<td>5.91</td>
<td>1.53</td>
</tr>
<tr>
<td>11</td>
<td>The package achieves its defined purpose</td>
<td>5.89</td>
<td>1.33</td>
</tr>
<tr>
<td>12</td>
<td>The user support materials are effective in assisting the user to run the program</td>
<td>5.72</td>
<td>1.53</td>
</tr>
<tr>
<td>13</td>
<td>Use of the package is motivational</td>
<td>5.65</td>
<td>1.39</td>
</tr>
<tr>
<td>Item importance priority</td>
<td>Evaluation criteria for educational game software programs</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------------------------------------------------------</td>
<td>------</td>
<td>-----</td>
</tr>
<tr>
<td>14</td>
<td>Feedback on student responses is effectively employed</td>
<td>5.65</td>
<td>1.53</td>
</tr>
<tr>
<td>15</td>
<td>The purpose of the package is well defined</td>
<td>5.63</td>
<td>1.42</td>
</tr>
<tr>
<td>16</td>
<td>The content is complete</td>
<td>5.52</td>
<td>1.84</td>
</tr>
<tr>
<td>17</td>
<td>Follows sound educational theory</td>
<td>5.52</td>
<td>1.92</td>
</tr>
<tr>
<td>18</td>
<td>The package effectively stimulates student creativity</td>
<td>5.48</td>
<td>1.43</td>
</tr>
<tr>
<td>19</td>
<td>The program encourages learning the subject materials</td>
<td>5.46</td>
<td>1.75</td>
</tr>
<tr>
<td>20</td>
<td>Follows sound educational techniques</td>
<td>5.37</td>
<td>1.79</td>
</tr>
<tr>
<td>21</td>
<td>Information displays are effective</td>
<td>5.37</td>
<td>1.88</td>
</tr>
<tr>
<td>22</td>
<td>The visual display is attractive, exciting, and absorbing</td>
<td>5.28</td>
<td>1.43</td>
</tr>
<tr>
<td>23</td>
<td>Amount of learning justifies time spent by users</td>
<td>5.16</td>
<td>2.02</td>
</tr>
<tr>
<td>24</td>
<td>The program appropriately uses relevant computer capabilities</td>
<td>5.07</td>
<td>2.15</td>
</tr>
<tr>
<td>25</td>
<td>The user support materials are comprehensive</td>
<td>4.89</td>
<td>1.90</td>
</tr>
<tr>
<td>26</td>
<td>The user can control the rate of presentation and review</td>
<td>4.89</td>
<td>2.06</td>
</tr>
<tr>
<td>27</td>
<td>Graphics is used for appropriate instructional reasons</td>
<td>4.74</td>
<td>1.77</td>
</tr>
<tr>
<td>28</td>
<td>The program uses a well-organized curriculum design</td>
<td>4.73</td>
<td>2.06</td>
</tr>
<tr>
<td>29</td>
<td>The system provides a summary of student progress</td>
<td>4.58</td>
<td>2.19</td>
</tr>
</tbody>
</table>
Table 17 indicates the item importance priority, mean, and standard deviation of evaluation criteria from secondary school teachers for drill/practice software programs.

Table 17 also indicates that nine evaluation criteria were rated above 6.00, 16 were rated from 5.00 to 5.99, six were rated from 4.00 to 4.99, and six were rated below 4.00.
TABLE 17. The item importance priority, mean, and standard deviation of evaluation criteria from secondary school teachers for drill/practice software programs

<table>
<thead>
<tr>
<th>Item importance priority</th>
<th>Evaluation criteria for drill/practice software programs</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The content is accurate</td>
<td>6.75</td>
<td>0.58</td>
</tr>
<tr>
<td>2</td>
<td>The program is reliable in normal use</td>
<td>6.56</td>
<td>0.74</td>
</tr>
<tr>
<td>3</td>
<td>The content has educational value</td>
<td>6.47</td>
<td>0.87</td>
</tr>
<tr>
<td>4</td>
<td>Presentation of content is clear, logical and well-organized</td>
<td>6.38</td>
<td>0.80</td>
</tr>
<tr>
<td>5</td>
<td>Teachers can easily employ the package</td>
<td>6.33</td>
<td>1.03</td>
</tr>
<tr>
<td>6</td>
<td>Intended users can independently operate the program</td>
<td>6.11</td>
<td>1.21</td>
</tr>
<tr>
<td>7</td>
<td>The content is free of race, ethnic, sex, and other stereotypes</td>
<td>6.09</td>
<td>1.70</td>
</tr>
<tr>
<td>8</td>
<td>The package achieves its defined purpose</td>
<td>6.09</td>
<td>1.02</td>
</tr>
<tr>
<td>9</td>
<td>The level of difficulty is appropriate for the target audience</td>
<td>6.06</td>
<td>0.90</td>
</tr>
<tr>
<td>10</td>
<td>The program encourages learning the subject material</td>
<td>5.96</td>
<td>1.26</td>
</tr>
<tr>
<td>11</td>
<td>Feedback on student responses is effectively employed</td>
<td>5.89</td>
<td>1.19</td>
</tr>
<tr>
<td>12</td>
<td>High student involvement and interaction</td>
<td>5.89</td>
<td>1.32</td>
</tr>
<tr>
<td>13</td>
<td>The purpose of the package is well defined</td>
<td>5.84</td>
<td>1.08</td>
</tr>
<tr>
<td>Item priority</td>
<td>Evaluation criteria for drill/practice software programs</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------------------------------------</td>
<td>------</td>
<td>----</td>
</tr>
<tr>
<td>14</td>
<td>The user support materials are effective in assisting the user to run the program</td>
<td>5.78</td>
<td>1.44</td>
</tr>
<tr>
<td>15</td>
<td>Amount of learning justifies time spent by users</td>
<td>5.73</td>
<td>1.60</td>
</tr>
<tr>
<td>16</td>
<td>The user can control the rate of presentation and review</td>
<td>5.73</td>
<td>1.60</td>
</tr>
<tr>
<td>17</td>
<td>Follows sound educational theory</td>
<td>5.67</td>
<td>1.21</td>
</tr>
<tr>
<td>18</td>
<td>The content is complete</td>
<td>5.63</td>
<td>1.78</td>
</tr>
<tr>
<td>19</td>
<td>Information displays are effective</td>
<td>5.61</td>
<td>1.11</td>
</tr>
<tr>
<td>20</td>
<td>Follows sound educational techniques</td>
<td>5.51</td>
<td>1.38</td>
</tr>
<tr>
<td>21</td>
<td>Use of the package is motivational</td>
<td>5.43</td>
<td>1.42</td>
</tr>
<tr>
<td>22</td>
<td>The user support materials are comprehensive</td>
<td>5.41</td>
<td>1.67</td>
</tr>
<tr>
<td>23</td>
<td>The system provides a summary of student progress</td>
<td>5.33</td>
<td>1.57</td>
</tr>
<tr>
<td>24</td>
<td>The program appropriately uses relevant computer capabilities</td>
<td>5.20</td>
<td>1.79</td>
</tr>
<tr>
<td>25</td>
<td>The program uses a well-organized curriculum design</td>
<td>5.16</td>
<td>1.78</td>
</tr>
<tr>
<td>26</td>
<td>Instruction is integrated with previous student experience</td>
<td>4.91</td>
<td>2.09</td>
</tr>
<tr>
<td>27</td>
<td>The user can control the sequence of presentation and review</td>
<td>4.80</td>
<td>2.00</td>
</tr>
<tr>
<td>28</td>
<td>The content is free of excessive violence</td>
<td>4.76</td>
<td>2.45</td>
</tr>
<tr>
<td>Item importance priority</td>
<td>Evaluation criteria for drill/practice software programs</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>--------------------------</td>
<td>--------------------------------------------------------</td>
<td>------</td>
<td>----</td>
</tr>
<tr>
<td>29</td>
<td>Student responses are automatically recorded for teacher use</td>
<td>4.67</td>
<td>2.15</td>
</tr>
<tr>
<td>30</td>
<td>Graphics is used for appropriate instructional reasons</td>
<td>4.40</td>
<td>1.73</td>
</tr>
<tr>
<td>31</td>
<td>The visual display is attractive, exciting, and absorbing</td>
<td>4.38</td>
<td>2.12</td>
</tr>
<tr>
<td>32</td>
<td>Learning is generalizable to an appropriate range of situations</td>
<td>3.95</td>
<td>2.08</td>
</tr>
<tr>
<td>33</td>
<td>The content of the program is available for inspection and/or change</td>
<td>3.91</td>
<td>2.46</td>
</tr>
<tr>
<td>34</td>
<td>The package effectively stimulates student creativity</td>
<td>3.42</td>
<td>2.58</td>
</tr>
<tr>
<td>35</td>
<td>The content is free of excessive competition</td>
<td>3.24</td>
<td>2.37</td>
</tr>
<tr>
<td>36</td>
<td>Color is used for appropriate instructional reasons</td>
<td>3.13</td>
<td>2.05</td>
</tr>
<tr>
<td>37</td>
<td>Sound is used for appropriate instructional reasons</td>
<td>3.05</td>
<td>2.13</td>
</tr>
</tbody>
</table>

Table 18 presents the item importance priority, mean, and standard deviation of evaluation criteria from secondary school teachers for educational game software programs.

Table 18 also indicated that five evaluation criteria were rated from 6.00 to 6.99, 11 were rated from 5.00 to 5.99, 13 were rated from 4.00 to 4.99, and eight were rated below 4.00.
TABLE 18. The item importance priority, mean, and standard deviation of evaluation criteria from secondary school teachers for educational game software programs

<table>
<thead>
<tr>
<th>Item importance priority</th>
<th>Evaluation criteria for educational game software programs</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The program is reliable in normal use</td>
<td>6.46</td>
<td>0.84</td>
</tr>
<tr>
<td>2</td>
<td>The content is accurate</td>
<td>6.39</td>
<td>1.36</td>
</tr>
<tr>
<td>3</td>
<td>Teachers can easily employ the package</td>
<td>6.04</td>
<td>1.42</td>
</tr>
<tr>
<td>4</td>
<td>Presentation of content is clear, logical, and well-organized</td>
<td>6.02</td>
<td>1.53</td>
</tr>
<tr>
<td>5</td>
<td>The content is free of race, ethnic, sex, and other stereotypes</td>
<td>6.00</td>
<td>1.86</td>
</tr>
<tr>
<td>6</td>
<td>Intended users can independently operate the program</td>
<td>5.98</td>
<td>1.23</td>
</tr>
<tr>
<td>7</td>
<td>High student involvement and interaction</td>
<td>5.96</td>
<td>1.37</td>
</tr>
<tr>
<td>8</td>
<td>The level of difficulty is appropriate for the target audience</td>
<td>5.90</td>
<td>1.03</td>
</tr>
<tr>
<td>9</td>
<td>The content had educational value</td>
<td>5.81</td>
<td>1.55</td>
</tr>
<tr>
<td>10</td>
<td>The package achieves its defined purpose</td>
<td>5.53</td>
<td>1.58</td>
</tr>
<tr>
<td>11</td>
<td>The user support materials are effective in assisting the user to run the program</td>
<td>5.46</td>
<td>1.82</td>
</tr>
<tr>
<td>12</td>
<td>Information displays are effective</td>
<td>5.36</td>
<td>1.68</td>
</tr>
<tr>
<td>13</td>
<td>Use of the package is motivational</td>
<td>5.23</td>
<td>1.56</td>
</tr>
<tr>
<td>14</td>
<td>The program appropriately uses relevant computer capabilities</td>
<td>5.15</td>
<td>1.80</td>
</tr>
<tr>
<td>Item importance priority</td>
<td>Evaluation criteria for educational game software programs</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------------------------------------------------------</td>
<td>------</td>
<td>-----</td>
</tr>
<tr>
<td>15</td>
<td>The purpose of the package is well defined</td>
<td>5.09</td>
<td>1.76</td>
</tr>
<tr>
<td>16</td>
<td>Follows sound educational theory</td>
<td>5.06</td>
<td>1.99</td>
</tr>
<tr>
<td>17</td>
<td>Feedback on student responses is effectively employed</td>
<td>4.89</td>
<td>2.02</td>
</tr>
<tr>
<td>18</td>
<td>The user support materials are comprehensive</td>
<td>4.87</td>
<td>2.11</td>
</tr>
<tr>
<td>19</td>
<td>Amount of learning justifies time spent by users</td>
<td>4.85</td>
<td>2.19</td>
</tr>
<tr>
<td>20</td>
<td>The user can control the rate of presentation and review</td>
<td>4.85</td>
<td>2.09</td>
</tr>
<tr>
<td>21</td>
<td>The content is complete</td>
<td>4.81</td>
<td>2.41</td>
</tr>
<tr>
<td>22</td>
<td>Follows sound educational techniques</td>
<td>4.76</td>
<td>2.08</td>
</tr>
<tr>
<td>23</td>
<td>The program encourages learning the subject material</td>
<td>4.76</td>
<td>2.30</td>
</tr>
<tr>
<td>24</td>
<td>The content is free of excessive violence</td>
<td>4.74</td>
<td>2.40</td>
</tr>
<tr>
<td>25</td>
<td>The visual display is attractive, exciting, and absorbing</td>
<td>4.70</td>
<td>2.04</td>
</tr>
<tr>
<td>26</td>
<td>Graphics is used for appropriate instructional reasons</td>
<td>4.52</td>
<td>1.94</td>
</tr>
<tr>
<td>27</td>
<td>The package effectively stimulates student creativity</td>
<td>4.15</td>
<td>2.46</td>
</tr>
<tr>
<td>28</td>
<td>The program uses a well-organized curriculum design</td>
<td>4.11</td>
<td>2.30</td>
</tr>
<tr>
<td>29</td>
<td>The system provides a summary of student progress</td>
<td>4.00</td>
<td>2.08</td>
</tr>
<tr>
<td>Item priority</td>
<td>Evaluation criteria for educational game software programs</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------------------------------------------</td>
<td>------</td>
<td>-----</td>
</tr>
<tr>
<td>30</td>
<td>The user can control the sequence of presentation and review</td>
<td>3.91</td>
<td>2.09</td>
</tr>
<tr>
<td>31</td>
<td>Instruction is integrated with previous student experience</td>
<td>3.65</td>
<td>2.25</td>
</tr>
<tr>
<td>32</td>
<td>Learning is generalizable to an appropriate range of situations</td>
<td>3.45</td>
<td>2.30</td>
</tr>
<tr>
<td>33</td>
<td>Color is used for appropriate instructional reasons</td>
<td>3.26</td>
<td>2.21</td>
</tr>
<tr>
<td>34</td>
<td>Sound is used for appropriate instructional reasons</td>
<td>3.21</td>
<td>2.24</td>
</tr>
<tr>
<td>35</td>
<td>Student responses are automatically recorded for teacher use</td>
<td>2.96</td>
<td>2.64</td>
</tr>
<tr>
<td>36</td>
<td>The content of the program is available for inspection and/or change</td>
<td>2.92</td>
<td>2.64</td>
</tr>
<tr>
<td>37</td>
<td>The content is free of excessive competition</td>
<td>2.38</td>
<td>2.44</td>
</tr>
</tbody>
</table>
TABLE 19. The frequency of the average score range of evaluation items for drill/practice software programs

<table>
<thead>
<tr>
<th></th>
<th>Combined-group</th>
<th>Elementary school teachers</th>
<th>Secondary school teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>6.00-6.99</td>
<td>11</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>5.00-5.99</td>
<td>15</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>4.00-4.99</td>
<td>7</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Below 4.00</td>
<td>4</td>
<td>1</td>
<td>6</td>
</tr>
</tbody>
</table>

TABLE 20. The frequency of the average score range of evaluation items for educational game software programs

<table>
<thead>
<tr>
<th></th>
<th>Combined-group</th>
<th>Elementary school teachers</th>
<th>Secondary school teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>6.00-6.99</td>
<td>8</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>5.00-5.99</td>
<td>15</td>
<td>16</td>
<td>11</td>
</tr>
<tr>
<td>4.00-4.99</td>
<td>7</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>Below 4.00</td>
<td>7</td>
<td>2</td>
<td>8</td>
</tr>
</tbody>
</table>
TABLE 21. Cronbach's alpha reliability coefficient for teachers and evaluation criteria qualities

<table>
<thead>
<tr>
<th></th>
<th>CQDP</th>
<th>CQEG</th>
<th>IQDP</th>
<th>IQEG</th>
<th>TQDP</th>
<th>TQEG</th>
</tr>
</thead>
<tbody>
<tr>
<td>ES</td>
<td>.6323</td>
<td>.5702</td>
<td>.8615</td>
<td>.8892</td>
<td>.7941</td>
<td>.7830</td>
</tr>
<tr>
<td>SS</td>
<td>.6291</td>
<td>.7274</td>
<td>.8318</td>
<td>.8849</td>
<td>.6947</td>
<td>.7404</td>
</tr>
</tbody>
</table>

CQDP : content quality of drill/practice software programs.
CQEG : content quality of educational game software programs.
IQDP : instructional quality of drill/practice software programs.
IQEG : instructional quality of educational game software programs.
TQDP : technical quality of drill/practice software programs.
TQEG : technical quality of educational game software programs.
ES : elementary school teachers.
SS : secondary school teachers.

TABLE 22. Spearman rank correlations based on item means between teacher groups and software programs

<table>
<thead>
<tr>
<th></th>
<th>DP/EG</th>
<th>ELE/SEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elementary school teachers</td>
<td>0.897</td>
<td>Drill/Practice 0.938</td>
</tr>
<tr>
<td>Secondary school teachers</td>
<td>0.901</td>
<td>Educational Game 0.920</td>
</tr>
</tbody>
</table>

DP/EG : comparison of the rank order between drill/practice and educational game software programs.
ELE/SEC : comparison of the rank order between elementary and secondary school teachers.
The means of evaluation criteria for drill/practice and educational game software programs are listed in Tables 19-20.

Table 21 presents Cronbach's alpha reliability for teachers and evaluation criteria qualities.

The Spearman rank correlation between teachers groups (elementary and secondary school teachers) and software programs (drill/practice and educational game) is presented in Table 22.
CHAPTER V. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The first four chapters of this study examined the background, related literature, methodology, and findings of the research undertaken. This chapter will summarize the study, and present conclusions and recommendations.

Restatement of the purpose

The purpose of this study was:

1. To identify which evaluation criteria will be necessary for the drill/practice and educational game software programs.

2. To identify for each evaluation criterion the average importance on evaluation guidelines in the process of evaluating drill/practice and educational game software programs.

3. To provide school teachers, administrators, and individuals with reference information for the evaluation of microcomputer software programs.

Restatement of the research questions

This study was designed to answer the following questions:

1. Is there any difference between the evaluation criteria for the drill/practice and educational game software programs?
2. Does each evaluation criterion gain the same importance (weighting score) in the process of evaluating drill/practice and educational game software programs?

3. If not, what will be the average importance (average scale value) for each evaluation criterion?

A review of related literature on evaluation criteria for microcomputer evaluation software programs was conducted and reported. The review of literature presented the following areas: overview, quantitative vs qualitative methods, evaluation criteria, weighting criteria, and summary.

A survey instrument (questionnaire) was designed with 37 questions on a likert seven-point scale. Additional information was collected regarding the respondent's teaching and computer background, the types of microcomputer software programs used in classroom teaching, and the situation in which microcomputer software is most useful. A pilot study was sent to 20 schools. After some modification, questionnaires were sent to a sample size of 301 schools which were randomly selected from a population of 1,506 elementary and secondary schools in the state of Iowa. The questionnaire was returned by 40% of the schools surveyed.
The findings of the chapter were based upon testing the hypotheses. A split-plot two-way analysis of variance used for testing hypotheses 1-3 revealed that there was a significant difference between the scale values of content quality evaluation criteria from elementary and secondary school teachers. This indicated that the two groups assigned different weighting scores to the evaluation criteria.

A split-plot two-way analysis of variance used for testing hypotheses 4-6 found there was a significant difference between scale values of content quality, instructional quality, and technical quality evaluation criteria from drill/practice and educational game software programs. This indicated that the two types of programs were assigned different weighting scores.

A split-plot two-way analysis of variance used for testing hypotheses 7-9 revealed that there was no interaction between scale values of content quality, instructional quality, and technical quality evaluation criteria from school teachers and software programs. This indicated that when different software programs were evaluated, they would not be affected by different groups.

A split-plot two-way analysis of variance used for answering research questions 1-3 found that there was a
difference between the evaluation criteria for the
drill/practice and educational game software programs. Most
of the evaluation criteria were not rated with the same
importance (weighting score) in the process of evaluating
drill/practice and educational game software programs. The
evaluation criterion scale values on importance were
reported by using the mean, standard deviation, and
priority. These values are presented in Tables 13-18.

Conclusions

The conclusions of this study will be presented in
terms of the stated hypotheses and research questions. Each
of them will be restated and followed by a conclusion and
discussion based upon the findings reported in Chapter Four.

Hypothesis 1

There was no significant difference between the scale
values of content quality evaluation criteria from responses
of elementary school teachers and secondary school teachers.

Conclusion of Hypothesis 1

Based upon the findings presented in Table 10, null
hypothesis 1 was rejected. There was a significant
difference between the scale values of content quality
evaluation criteria from elementary school teachers and
secondary school teachers.
Discussion of Hypothesis 1

Elementary school teachers and secondary school teachers assigned different weighting scores to the content quality evaluation criteria. They thought that when drill/practice and education game software programs are evaluated, different weighting scores should be given to content quality evaluation criteria. Traditional equal-weighting-score approach is not recommended.

Hypothesis 2

There was no significant difference between the scale values of instructional quality evaluation criteria from responses of elementary school teachers and secondary school teachers.

Conclusion of Hypothesis 2

It was concluded, based upon findings reported in Table 11, that there was no significant difference between the scale values of instructional quality evaluation criteria from elementary school teachers and secondary school teachers. Therefore, the null hypothesis was not rejected.

Discussion of Hypothesis 2

From the above conclusion, elementary and secondary school teachers considered that same weighting scores should be given to drill/practice and educational game software programs for instructional quality evaluation criteria.
Hypothesis 3

There was no significant difference between the scale values of technical quality evaluation criteria from responses of elementary school teachers and secondary school teachers.

Conclusion of Hypothesis 3

Based upon findings presented in Table 12, null Hypothesis 3 was not rejected. There was no significant difference between the scale values of technical quality from elementary school teachers and secondary school teachers.

Discussion of Hypothesis 3

Just like the previous discussion, elementary and secondary school teachers also recommended that same weighting scores should be applied to drill/practice and educational game software programs for technical quality evaluation criteria.

Hypothesis 4

There was no significant difference between the scale values of content quality evaluation criteria from drill/practice and educational game software programs.
Conclusion of Hypothesis 4

Based upon the findings presented in Table 10, null Hypothesis 4 was rejected. There was a significant difference between the scale values of content quality evaluation criteria from drill/practice and educational game software programs.

Discussion of Hypothesis 4

From the above conclusion, in the process of evaluating drill/practice and educational game software programs, different weighting scores would be assigned to content quality evaluation criteria. Same weighting scores for different types of programs are not suggested.

Hypothesis 5

There was no significant difference between the scale values of instructional quality evaluation criteria from drill/practice and educational game software programs.

Conclusion of Hypothesis 5

It can be concluded, based upon the findings presented in Table 11, that there was a significant difference between the scale values of instructional quality evaluation criteria from drill/practice and educational game software programs.
Discussion of Hypothesis 5

For instructional quality evaluation criteria, different weighting scores were recommended to be assigned to drill/practice and educational game software programs during the process of evaluating software programs.

Hypothesis 6

There was no significant difference between the scale values of technical quality evaluation criteria from drill/practice and educational game software programs.

Conclusion of Hypothesis 6

It was concluded, based upon findings reported in Table 12, that there was a significant difference between the scale values of technical quality evaluation criteria from drill/practice and educational game software programs.

Discussion of Hypothesis 6

It is also recommended that during the process of evaluating drill/practice and educational game software programs using technical quality evaluation criteria, different weighting scores should be assigned to those two types of programs.
**Hypothesis 7**

There was no interaction between the scale values of content quality evaluation criteria from school teachers and software programs.

**Conclusion of Hypothesis 7**

Based upon the findings presented in Table 10, null hypothesis 7 was not rejected. There was no interaction between the scale values of content quality evaluation criteria from school teachers and software programs.

**Discussion of Hypothesis 7**

During the evaluating process of microcomputer software programs, elementary and secondary school teachers' ratings would not be affected by the different types of software programs using content quality evaluation criteria. This means that school teachers would give weighting scores independently to drill/practice and educational game software programs.

**Hypothesis 8**

There was no interaction between the scale values of instructional quality evaluation criteria from school teachers and software programs.
Conclusion of Hypothesis 8

Based upon the findings presented in Table 11, null hypothesis 8 was not rejected. There was no interaction between the scale values of instructional quality evaluation criteria from school teachers and software programs.

Discussion of Hypothesis 8

When elementary and secondary school teachers use instructional quality evaluation criteria to evaluate drill/practice and educational game software programs, they will assign weighting score to each criterion independently, no matter what types of software programs they evaluate.

Hypothesis 9

There was no interaction between the scale values of technical quality evaluation criteria from school teachers and software programs.

Conclusion of Hypothesis 9

It was concluded, based upon findings reported in Table 12, that there was no interaction between the scale values of technical quality evaluation criteria from school teachers and software programs.
Discussion of Hypothesis 9

Each weighting score will be assigned independently to each technical evaluation criterion by elementary and secondary school teachers during the process of evaluating drill/practice and educational game software programs.

Research Questions

1. Is there any difference between the evaluation criteria for the drill/practice and educational game software programs?

2. Does each evaluation criterion gain the same importance (weighting score) in the process of evaluating drill/practice and educational game software programs?

3. If not, what will be the average importance (average scale value) for each evaluation criterion?

Conclusion of Research Questions

Based upon the findings of Chapter Four, there was a difference between the evaluation criteria for the drill/practice and educational game software programs. Few evaluation criteria gained the same weighting score. Each evaluation criterion average importance and its priority was presented on Tables 13-18.
Discussion of Research Questions

Obviously, there did exist differences between the scale values of drill/practice and educational game software programs. When different software programs are to be evaluated, it is suggested that different sets of evaluation criteria should be applied. In addition, different weighting scores were also suggested for the evaluation criteria. When elementary and secondary school teachers evaluate drill/practice and educational game software programs, they can make reference to Tables 13-18, which may help them to make better decisions about the quality of software programs.

Table 19 revealed that the combined-group rated most (70%) of the evaluation criteria items of high average weighting scores (5.00-6.99) to drill/practice software programs. Elementary school teachers rated as high as 81% of the evaluation criteria items of average weighting scores (5.00-6.99) to drill/practice software programs. On the other hand, secondary school teachers rated 68% of the evaluation items of average weighting scores (5.00-6.99) to drill/practice software programs.

Table 20 reveals that the combined-group gave 62% of the evaluation criteria items of average weighting scores (5.00-6.99) to educational game software programs.
Elementary school teachers rated 65% of the average weighting scores (5.00-6.99) to educational game software programs. On the other hand, secondary school teachers rated 43% of the evaluation criteria items of average weighting scores (5.00-6.99) to educational game software programs.

Generally speaking, above one-half of the evaluation criteria items were thought to be very important by elementary and secondary school teachers.

Table 21 presents the internal reliability when teachers evaluate different qualities of evaluation criteria. A fairly high reliability coefficient can be observed from the table. The CQDP and CQEG contain eight items of evaluation criteria respectively. The IQDP and IQEG include 19 items of evaluation criteria respectively. Both the TQDP and TQEG include ten items of evaluation criteria respectively. When the item number increases, the reliability coefficient increases also. This shows a reasonable trend as expected.

From Table 22, high spearman rank correlation between drill/practice and educational game software programs can be found when elementary school teachers and secondary school teachers evaluated those programs respectively. Also, high spearman rank correlation between elementary and secondary
school teachers can be found when drill/practice and educational game software programs were evaluated respectively.

Recommendations

The recommendations of this study are based upon the findings and conclusions that were presented. The recommendations are presented in two parts. The first part contains recommendations based on the findings and conclusions; the second part presents recommendations for further research related to this study.

Recommendations based on the findings and conclusions

1. Any evaluation guidelines should take into account variations in instructional software. A uniform set of evaluation criteria are not recommended for application to all software materials.

2. Each evaluation criterion may contribute different effectiveness to the program. The weighting score of each criterion presented in Chapter Four can be applied carefully to the evaluation guidelines in the process of evaluating the quality of software programs.
3. For most of the evaluation criteria items (instructional quality and technical quality) the weighting scores assigned by elementary and secondary school teachers did not differ significantly. Therefore, the same set of evaluation criteria can be applied by these two groups of evaluators.

4. Individuals who are interested in software evaluations may refer to Tables 13-14 for the evaluation of drill/practice and educational game software programs.

5. The different sets of evaluation criteria for drill/practice and educational game software programs in this study can be modified for specific needs.

6. The results of this study were subjective in nature. Therefore, any application from this study may contain a certain degree variance and bias. These criteria should be carefully applied to the evaluation of software programs.

**Recommendations for further study**

1. Considering the time the teachers spent in responding to the questionnaire, this study only
covered two types of software programs. Investigation of other types of software may be continued in future studies.

2. Given available time and funds, the researcher suggests that other states (not only the state of Iowa) be chosen for replication studies. This could make the results of the study more generalizable.

3. If funds and time are available, the researcher recommends that more external data, especially student achievement gain scores which related to a specific evaluation criterion, should be collected to establish the validity of the evaluation criteria.
BIBLIOGRAPHY


Blease, Derek. (1986). Evaluating educational software. Dover, NH: Croom Helm Ltd.


I wish to gratefully acknowledge the support, time, and energy of many individuals who have made it possible for me to complete my doctoral program at Iowa State University during the past years.

First, thank you to my program of study committee members, Dr. Trevor Howe, Dr. William Miller, Dr. William Wolansky, Dr. Richard Warren, and Dr. Arthur Akers for their guidance and support throughout this program. Also, thank you to Dr. Federick Graham, Dr. Tony Netusil, and Professor Albert Sherick for serving on the committee at the final examination. In addition, thank you Dr. Leroy Wolins for helping me to solve SAS computer programming and statistical problems.

To Dr. Trevor Howe, my major professor, I exceedingly appreciate the countless meetings and close assistance he gave me throughout my program. Without his assistance and encouragement, this dissertation would not have been completed.

Finally, to my parents, Song-Tau and Tzei-Shung, my deepest gratitude for their sacrifice and financial support. Without their support, my doctoral program would not have been completed.
APPENDIX A. LETTER ACCOMPANYING PILOT INSTRUMENT
April 17, 1987

Dear Principal:

As the use of microcomputers proliferate in the educational setting, the production of suitable instructional materials for the microcomputer has become an issue of great importance and concern to educators. There is a great need to determine what exactly is "good" instructional software and what is inadequate software. As you know, adequate evaluation criteria are very important for the evaluation of instructional software programs.

We need your school's input into this study to identify evaluation criteria items and item importance for the drill/practice and educational game software programs. Enclosed is one questionnaire for the teacher who has the experience of using drill/practice and/or educational game software programs in his/her classroom teaching. Such experience will help us much in the success of this study.

Your assistance in forwarding the questionnaire to the appropriate teacher is very important for the success of this study.

Your time and assistance are very much appreciated. Thank you very much.

Sincerely,

Yoau-Chau Jeng
Doctoral Student
Department of Industrial Education & Technology

Trevor G. Howe
Professor & Chairperson
Department of Industrial Education & Technology
Dear Teacher:

As the use of microcomputers proliferate in the educational setting, the production of suitable instructional materials for the microcomputer has become an issue of great importance and concern to educators. There is a great need to determine what exactly is "good" instructional software and what is inadequate software. As you know, adequate evaluation criteria are very important for the evaluation of instructional software programs.

We need your input into this study to identify evaluation criteria items and item importance for the drill/practice and educational game software programs.

All data will be kept strictly confidential. The number that appears on the survey questionnaire is for the purpose of a follow-up letter to non-respondents. All information will be collected and summarized as a group thus protecting your confidentiality. The survey will require about 15 minutes of your time to complete.

After you have completed the questionnaire, please deposit it in the U.S. mail. No postage is necessary. Please have your questionnaire completed and returned to us within next 7 days.

If you are interested in the results, you may receive a summary of them by writing "COPY OF RESULTS REQUESTED" and your name and address on the back of the enclosed return envelope. Please do not put this information on the questionnaire form itself.

Your time and cooperation are very much appreciated. Thank you very much.

Sincerely,

Yoau-Chau Jeng
Doctoral Student
Department of Industrial Education & Technology

Trevor G. Howe
Professor & Chairperson
Department of Industrial Education & Technology
APPENDIX B. PILOT SURVEY INSTRUMENT
PART A: BASIC DATA

1. Gender: __ Male __ Female

2. Which grade are you teaching in your school:
   __ K-1 __ 2 __ 3 __ 4 __ 5 __ 6
   __ 7 __ 8 __ 9 __ 10 __ 11 __ 12

3. Type(s) of microcomputer software program(s) you use in your classes:
   __ Drill/Practice __ Educational Game __ Tutorial
   __ Simulation __ Problem Solving __ Management/Administration
   __ Informational __ Testing __ Word Processing __ Utility
   __ Other, please specify ________________________

4. Number of years of teaching experience:
   __ 1-5 __ 6-10 __ 11-15 __ 16-20 __ More than 20

5. Number of years experience with computers:
   __ 1-5 __ 6-10 __ 11-15 __ 16-20 __ More than 20

6. Experience of using drill/practice software programs:
   __ Yes __ No

7. Experience of using educational game software programs:
   __ Yes __ No

8. Computer software is most useful as a:
   __ Substitute for other method instruction
   __ Supplement to other method instruction
   __ Means of requiring needed computer skills
   __ Currently isn't very useful
**PART B: EVALUATION CRITERIA**

**DIRECTIONS:** The following is a list of evaluation criteria for microcomputer software programs. You are asked to verify the necessity of the criteria and to rate the importance (weighting score) of each criterion for the Drill/Practice (DP) and Educational Game (EG) software programs respectively. If you think that criterion is not necessary for the software evaluation, please check Not Necessary (NN). Then check the next criterion. If you check Necessary (N) for that criterion, please rate the importance by using scale value which runs from 1 to 7 with the 1 being little important and the 7 being absolutely important. If you think of any evaluation criterion which is not listed below, please specify, then rate the importance for that criterion. Please indicate the appropriate scale value by circling the respective number which was listed below:

<table>
<thead>
<tr>
<th>Scale Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Little Important</td>
</tr>
<tr>
<td>2</td>
<td>Somewhat Important</td>
</tr>
<tr>
<td>3</td>
<td>Moderately Important</td>
</tr>
<tr>
<td>4</td>
<td>Important</td>
</tr>
<tr>
<td>5</td>
<td>Very Important</td>
</tr>
<tr>
<td>6</td>
<td>Strongly Important</td>
</tr>
<tr>
<td>7</td>
<td>Absolutely Important</td>
</tr>
</tbody>
</table>

**EVALUATION CRITERIA**

* DP: Drill/Practice  
* EG: Educational Game

**Characteristics: Content Quality**

1. The content is accurate.  
   - DP:  
   - EG: 
   - (NN) N 1 2 3 4 5 6 7

2. The content is complete.  
   - DP:  
   - EG: 
   - (NN) N 1 2 3 4 5 6 7

3. The content has educational value.  
   - DP:  
   - EG: 
   - (NN) N 1 2 3 4 5 6 7

4. The content is free of excessive competition or violence.  
   - DP:  
   - EG: 
   - (NN) N 1 2 3 4 5 6 7

5. The content is free of race, ethnic, sex, and other stereotypes.  
   - DP:  
   - EG: 
   - (NN) N 1 2 3 4 5 6 7

6. Amount of learning justifies time spent by users.  
   - DP:  
   - EG: 
   - (NN) N 1 2 3 4 5 6 7

7. The content of the program is available for inspection and/or change.  
   - DP:  
   - EG: 
   - (NN) N 1 2 3 4 5 6 7
Characteristics: Content Quality

Other evaluation criteria, please specify:

8. .......................................................... (DP) ...... 1 2 3 4 5 6 7
    .......................................................... (EG) ...... 1 2 3 4 5 6 7

9. .......................................................... (DP) ...... 1 2 3 4 5 6 7
    .......................................................... (EG) ...... 1 2 3 4 5 6 7

Characteristics: Instructional Quality

1. The purpose of the package is well defined.......................... (DP) ______ 1 2 3 4 5 6 7
    .......................................................... (EG) ______ 1 2 3 4 5 6 7

2. The package achieves its defined purpose................................ (DP) ...... 1 2 3 4 5 6 7
    .......................................................... (EG) ...... 1 2 3 4 5 6 7

3. Presentation of content is clear, logical and well-organized........ (DP) ______ 1 2 3 4 5 6 7
    .......................................................... (EG) ______ 1 2 3 4 5 6 7

4. The level of difficulty is appropriate for the target audience........... (DP) ______ 1 2 3 4 5 6 7
    .......................................................... (EG) ______ 1 2 3 4 5 6 7

5. Graphics/color/sound are used for appropriate instructional reasons..... (DP) ______ 1 2 3 4 5 6 7
    .......................................................... (EG) ______ 1 2 3 4 5 6 7

6. Use of the package is motivational........................................ (DP) ______ 1 2 3 4 5 6 7
    .......................................................... (EG) ______ 1 2 3 4 5 6 7

7. The package effectively stimulates student creativity..................... (DP) ______ 1 2 3 4 5 6 7
    .......................................................... (EG) ______ 1 2 3 4 5 6 7

8. Feedback on student responses is effectively employed.................... (DP) ______ 1 2 3 4 5 6 7
    .......................................................... (EG) ______ 1 2 3 4 5 6 7

9. The user can control the rate and sequence of presentation and review..... (DP) ______ 1 2 3 4 5 6 7
    .......................................................... (EG) ______ 1 2 3 4 5 6 7

10. Instruction is integrated with previous student experience.............. (DP) ______ 1 2 3 4 5 6 7
    .......................................................... (EG) ______ 1 2 3 4 5 6 7
11. Learning is generalizable to an appropriate range of situations........ (DP) ______ 1 2 3 4 5 6 7
........ (EG) ______ 1 2 3 4 5 6 7

12. Well-organized curriculum design.............. (DP) ______ 1 2 3 4 5 6 7
........ (EG) ______ 1 2 3 4 5 6 7

13. Follows sound educational theory............. (DP) ______ 1 2 3 4 5 6 7
........ (EG) ______ 1 2 3 4 5 6 7

14. Follows sound educational techniques........ (DP) ______ 1 2 3 4 5 6 7
........ (EG) ______ 1 2 3 4 5 6 7

15. The program encourages cooperation.......... (DP) ______ 1 2 3 4 5 6 7
........ (EG) ______ 1 2 3 4 5 6 7

16. High student involvement and interaction........................ (DP) ______ 1 2 3 4 5 6 7
........................ (EG) ______ 1 2 3 4 5 6 7

Other evaluation criteria, please specify:

17. .................................................. (DP) ....... 1 2 3 4 5 6 7
.................................................. (EG) ....... 1 2 3 4 5 6 7

18. .................................................. (DP) ....... 1 2 3 4 5 6 7
.................................................. (EG) ....... 1 2 3 4 5 6 7

Characteristics: Technical Quality
1. The user support materials are comprehensive........ (DP) ______ 1 2 3 4 5 6 7
........ (EG) ______ 1 2 3 4 5 6 7

2. The user support materials are effective............. (DP) ______ 1 2 3 4 5 6 7
........ (EG) ______ 1 2 3 4 5 6 7

3. Information displays are effective............. (DP) ______ 1 2 3 4 5 6 7
........ (EG) ______ 1 2 3 4 5 6 7

4. Intended users can and independently operate the program........ (DP) ______ 1 2 3 4 5 6 7
........ (EG) ______ 1 2 3 4 5 6 7

5. Teachers can easily employ the package............. (DP) ______ 1 2 3 4 5 6 7
........ (EG) ______ 1 2 3 4 5 6 7
Characteristics: Technical Quality

6. The program is reliable in normal use... (DP) ___ ___ 1 2 3 4 5 6 7
   .... (EG) ___ ___ 1 2 3 4 5 6 7

7. The program appropriately uses relevant computer capabilities......... (DP) ___ ___ 1 2 3 4 5 6 7
   ........... (EG) ___ ___ 1 2 3 4 5 6 7

8. The visual display is attractive, exciting, and absorbing................ (DP) ___ ___ 1 2 3 4 5 6 7
   ............... (EG) ___ ___ 1 2 3 4 5 6 7

Other evaluation criteria, please specify:

9. ........................................................................... (DP) ...... 1 2 3 4 5 6 7
    ........................................................................... (EG) ...... 1 2 3 4 5 6 7

10. ........................................................................... (DP) ...... 1 2 3 4 5 6 7
    ....................................................................... (EG) ...... 1 2 3 4 5 6 7

PART C: please complete the information below about the survey you have just completed.

1. Are directions clear enough to let you know how to respond the question?
   __ Yes   __ No

2. How many minutes did it take you to complete parts A and B? ______

3. What questions, if any, were confusing to answer? ________________

4. Please list other comments you may have:
APPENDIX C. LETTER ACCOMPANYING SURVEY INSTRUMENT
Dear Principal:

As the use of microcomputers proliferate in the educational setting, the production of suitable instructional materials for the microcomputer has become an issue of great importance and concern to educators. There is a great need to determine what exactly is "good" instructional software and what is inadequate software. As you know, adequate evaluation criteria are very important for the evaluation of instructional software programs.

We need your school's input into this study to identify evaluation criteria items and item importance for the drill/practice and educational game software programs. Enclosed is one questionnaire for the teacher who has the experience of using drill/practice and/or educational game software programs in his/her classroom teaching. Such experience will help us much in the success of this study.

Your assistance in forwarding the questionnaire to the appropriate teacher is very important for the success of this study.

Your time and assistance are very much appreciated. Thank you very much.

Sincerely,

Yoau-Chau Jeng
Doctoral Student
Department of Industrial Education & Technology

Trevor G. Howe
Professor & Chairperson
Department of Industrial Education & Technology
Dear Teacher:

As the use of microcomputers proliferate in the educational setting, the production of suitable instructional materials for the microcomputer has become an issue of great importance and concern to educators. There is a great need to determine what exactly is "good" instructional software and what is inadequate software. As you know, adequate evaluation criteria are very important for the evaluation of instructional software programs.

We need your input into this study to identify evaluation criteria items and item importance for the drill/practice and educational game software programs.

All data will be kept strictly confidential. The number that appears on the survey questionnaire is for the purpose of a follow-up letter to nonrespondents. All information will be collected and summarized as a group thus protecting your confidentiality. The survey will require about 15 minutes of your time to complete.

After you have completed the questionnaire, please deposit it in the U.S. mail. No postage is necessary. Please have your questionnaire completed and returned to us by May 18, 1987.

If you are interested in the results, you may receive a summary of them by writing "COPY OF RESULTS REQUESTED" and your name and address on the back of the enclosed return envelope. Please do not put this information on the questionnaire form itself.

Your time and cooperation are very much appreciated. Thank you very much.

Sincerely,

Yoau-Chau Jeng
Doctoral Student
Department of Industrial Education & Technology

Trevor G. Howe
Professor & Chairperson
Department of Industrial Education & Technology
APPENDIX D. SURVEY INSTRUMENT WITH ADDENDUM
PART A: BASIC DATA

1. Gender: __Male  __Female

2. Which grade are you teaching in your school:
   ___K-1  ___2  ___3  ___4  ___5  ___6
   ___7  ___8  ___9  ___10  ___11  ___12

3. Type(s) of microcomputer software program(s) you use in your classes:
   ___Drill/Practice  ___Educational Game  ___Tutorial
   ___Simulation  ___Problem Solving  ___Management/Administration
   ___Informational  ___Testing  ___Word Processing
   ___Utility  ___Other, please specify_________________________

4. Number of years of teaching experience:
   ___1-5  ___6-10  ___11-15  ___16-20  ___More than 20

5. Number of years experience with computers:
   ___1-5  ___6-10  ___11-15  ___16-20  ___More than 20

6. Experience of using drill/practice software programs:
   ___Yes  ___No

7. Experience of using educational game software programs:
   ___Yes  ___No

8. Computer software is most useful as a:
   ___Substitute for other methods of instruction
   ___Supplement to other methods of instruction
   ___Means for developing computer skills
   ___Currently isn't very useful
PART B: EVALUATION CRITERIA

DIRECTIONS: The following is a list of evaluation criteria for microcomputer software programs. You are asked to verify the necessity of the criteria and to rate the importance (weighting score) of each criterion for the Drill/Practice (DP) and Educational Game (EG) software programs respectively. If you think that criterion is not necessary for the software evaluation, please check Not Necessary (NN). Then check the next criterion. If you check Necessary (N) for that criterion, please rate the importance by using scale value which runs from 1 to 7 with the 1 being little important and the 7 being absolutely important. Please indicate the appropriate scale value by circling the respective number which was listed below:

Little Important .............. 1
Somewhat Important ............. 2
Moderately Important ........... 3
Important ....................... 4
Very Important ................. 5
Strongly Important ............. 6
Absolutely Important ........... 7

<table>
<thead>
<tr>
<th>EVALUATION CRITERIA</th>
<th>Characteristics: Content Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NN  N  LI  AI</td>
</tr>
<tr>
<td>1. The content is accurate.</td>
<td>(DP) 1 2 3 4 5 6 7</td>
</tr>
<tr>
<td></td>
<td>(EG) 1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>2. The content is complete.</td>
<td>(DP) 1 2 3 4 5 6 7</td>
</tr>
<tr>
<td></td>
<td>(EG) 1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>3. The content has educational value.</td>
<td>(DP) 1 2 3 4 5 6 7</td>
</tr>
<tr>
<td></td>
<td>(EG) 1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>4. The content is free of excessive competition.</td>
<td>(DP) 1 2 3 4 5 6 7</td>
</tr>
<tr>
<td></td>
<td>(EG) 1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>5. The content is free of excessive violence.</td>
<td>(DP) 1 2 3 4 5 6 7</td>
</tr>
<tr>
<td></td>
<td>(EG) 1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>6. The content is free of race, ethnic, sex, and other stereotypes.</td>
<td>(DP) 1 2 3 4 5 6 7</td>
</tr>
<tr>
<td></td>
<td>(EG) 1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>7. Amount of learning justifies time spent by users.</td>
<td>(DP) 1 2 3 4 5 6 7</td>
</tr>
<tr>
<td></td>
<td>(EG) 1 2 3 4 5 6 7</td>
</tr>
</tbody>
</table>
Characteristics: Content Quality

8. The content of the program is available for inspection and/or change. 

Characteristics: Instructional Quality

1. The purpose of the package is well defined. 

2. The package achieves its defined purpose. 

3. Presentation of content is clear, logical and well-organized. 

4. The level of difficulty is appropriate for the target audience. 

5. Graphics is used for appropriate instructional reasons. 

6. Color is used for appropriate instructional reasons. 

7. Sound is used for appropriate instructional reasons. 

8. Use of the package is motivational. 

9. The package effectively stimulates student creativity. 

10. Feedback on student responses is effectively employed. 

11. The user can control the rate of presentation and review.
Characteristics: Instructional Quality

12. The user can control the sequence of presentation and review. (DP) 1 2 3 4 5 6 7 (EG) 1 2 3 4 5 6 7

13. Instruction is integrated with previous student experience. (DP) 1 2 3 4 5 6 7 (EG) 1 2 3 4 5 6 7

14. Learning is generalizable to an appropriate range of situations. (DP) 1 2 3 4 5 6 7 (EG) 1 2 3 4 5 6 7

15. The program uses a well-organized curriculum design. (DP) 1 2 3 4 5 6 7 (EG) 1 2 3 4 5 6 7

16. Follows sound educational theory. (DP) 1 2 3 4 5 6 7 (EG) 1 2 3 4 5 6 7

17. Follows sound educational techniques. (DP) 1 2 3 4 5 6 7 (EG) 1 2 3 4 5 6 7

18. The program encourages learning the subject material. (DP) 1 2 3 4 5 6 7 (EG) 1 2 3 4 5 6 7

19. High student involvement and interaction. (DP) 1 2 3 4 5 6 7 (EG) 1 2 3 4 5 6 7

Characteristics: Technical Quality

1. The user support materials are comprehensive. (DP) 1 2 3 4 5 6 7 (EG) 1 2 3 4 5 6 7

2. The user support materials are effective in assisting the user to run the program. (DP) 1 2 3 4 5 6 7 (EG) 1 2 3 4 5 6 7

3. Information displays are effective. (DP) 1 2 3 4 5 6 7 (EG) 1 2 3 4 5 6 7

4. Intended users can independently operate the program. (DP) 1 2 3 4 5 6 7 (EG) 1 2 3 4 5 6 7
Characteristics: Technical Quality

5. Teachers can easily employ the package. ......................(DP) 1 2 3 4 5 6 7

6. The program is reliable in normal use. ......................(DP) 1 2 3 4 5 6 7

7. The program appropriately uses relevant computer capabilities. ....(DP) 1 2 3 4 5 6 7

8. The visual display is attractive, exciting, and absorbing. .........(DP) 1 2 3 4 5 6 7

9. The system provides a summary of student progress. ..............(DP) 1 2 3 4 5 6 7

10. Student responses are automatically recorded for teacher use. .......(DP) 1 2 3 4 5 6 7
Addendum

Please note the following information for your use:

There are many ways to classify microcomputer software programs. We are narrowing this study to only two classifications from (Part A, Basic Data, Number 3). We have selected Drill/Practice and Educational Game. For the remainder of the questionnaire, as you respond, a strong point to remember is to distinguish between Drill/Practice and Educational Game.

To further aid you, the following definitions and examples of each are described.

**Drill/Practice:** this is the most commonly used software in education. Its purpose is to reinforce basic skill, a task the computer does extremely well. These programs provide the student with a number of basic questions generated by the computer itself or by the teacher. Drill/Practice programs have the ability to tally a student’s correct and incorrect responses, and frequently will report to the teacher the number of correct and incorrect answers. They may also recommend and provide extra practice if this is appropriate. Students receive feedback from Drill/Practice programs. This can be as simple as yes, great, or O.K., or an elaborate pictorial response with a lot of movement on the screen. SRA Computer Drill and Instruction, Fact Track, Milliken Math, Estimation, and Comprehension Power are good examples of Drill/Practice software programs.

**Educational Game:** Educational Game is designed primarily for instructional purposes. Generally, it teaches or reinforces an academic or manual skill (typing). It is usually pleasurable; but fun is a secondary consideration. The student may or may not have control over timing and sequence of events, depending upon the educational purposes. The emphasis is on using intellectual skills rather than developing strategies or solutions. The reinforcement for learning is immediate. Computer graphics and sound are included to enhance learning. The game is nontargeting to the player. The player experiences successful learning when performance is effective. The educational game tends to be less intense than an arcade game. Master Type, Hangman, Robot War, and Snooper Troops are good examples of educational game software programs.
APPENDIX E. FOLLOW-UP LETTER
May 21, 1987

Dear Principal:

Two weeks ago a questionnaire was mailed to you asking you to forward to the teacher who is experienced with the Drill/Practice and Educational Game software programs.

Please forward this follow-up letter and enclosed card to this teacher. The involvement of your teacher is very important in the success of this study.

Your assistance is very much appreciated during this semester-end period. Thank you very much.

Sincerely,

Yoau-Chau Jeng
Doctoral Student
Department of Industrial Education & Technology

Trevor G. Howe, PhD
Professor and Chairperson
Department of Industrial Education & Technology

enclosure
May 21, 1987

Dear Teacher:

Two weeks ago a questionnaire was mailed to your principal, then supposedly forwarded to you, asking your input into identifying the evaluation criteria and item importance for the Drill/Practice and Educational Game software programs.

If you have completed the questionnaire, please mail it to us as soon as possible. If not, please have your questionnaire completed and mailed to us by June 1. You were randomly selected for this study and therefore your response is crucial for a complete and accurate study.

To encourage your cooperation, a chance (1:10) to win a gift certificate to Bonanza (worth $6.00) or McDonald's (worth $4.00) will make you eligible if you have returned it to us by June 1, 1987.

If you need another questionnaire, please do not hesitate to complete the green card and mail it to us. No postage is necessary. We will immediately mail you a new questionnaire.

During this semester-end period your time and cooperation will be very much appreciated. Thank you very much.

Sincerely,

Yoau-Chau Jeng
Doctoral Student
Department of Industrial Education & Technology

Trevor G. Howe, PhD
Professor and Chairperson
Department of Industrial Education & Technology

enclosure
APPENDIX F. HUMAN SUBJECTS IN RESEARCH CONSENT FORM
INFORMATION ON THE USE OF HUMAN SUBJECTS IN RESEARCH
IOWA STATE UNIVERSITY

(Please follow the accompanying instructions for completing this form.)

1. Title of project (please type): The identification of evaluation criteria
   items and item importance of drill/practice and educational game software programs

2. I agree to provide the proper surveillance of this project to insure that the rights
   and welfare of the human subjects are properly protected. Additions to or changes
   in procedures affecting the subjects after the project has been approved will be
   submitted to the committee for review.

   Jennt. Yoau-chau
   Typed Name of Principal Investigator
   Date

   Room B-7, I. ED. II
   Campus Address
   204-8520
   Campus Telephone

3. Signatures of others (If any) Date Relationship to Principal Investigator

4. ATTACH an additional page(s) (A) describing your proposed research and (B) the
   subjects to be used, (C) indicating any risks or discomforts to the subjects, and
   (D) covering any topics checked below. CHECK all boxes applicable.
   □ Medical clearance necessary before subjects can participate
   □ Samples (blood, tissue, etc.) from subjects
   □ Administration of substances (foods, drugs, etc.) to subjects
   □ Physical exercise or conditioning for subjects
   □ Deception of subjects
   □ Subjects under 14 years of age and/or □ Subjects 14-17 years of
   □ Subjects in institutions
   □ Research must be approved by another institution or agency

5. ATTACH an example of the material to be used to obtain informed consent and CHECK
   which type will be used.
   □ Signed informed consent will be obtained.
   □ Modified informed consent will be obtained.

6. Anticipated date on which subjects will be first contacted: 4 15 87
   Anticipated date for last contact with subjects: 6 01 87

7. If Applicable: Anticipated date on which audio or visual tapes will be erased and/or
   identifiers will be removed from completed survey instruments: 6 15 87

8. Signature of Head or Chairperson Date Department or Administrative Unit

9. Decision of the University Committee on the Use of Human Subjects in Research:
   □ Project Approved □ Project not approved □ No action required

   Name of Committee Chairperson Date Signature of Committee Chairperson