Contents

ARTICLES

A Comparison of the Leadership Behaviors of Instructional Designers in Higher Education and Industry, Barry Willis 2

Issues in Television-Centered Instruction for Adults, Penelope L. Richardson 6

Measuring the Importance of Learning from Instruction: The Outcome-Consequence Model of Evaluation, Michael J. Hannafin 14

Notetaking Research, Carol A. Carrier 19

The Number of Performance Assessments Necessary to Determine Competence, Robert L. Lathrop 26

DEPARTMENTS

Book Reviews, Allison Rossett 32


Who's Hiring Who, by Richard Lathrop

The Professional Job Changing System, by Robert Jameson Gerberg

Sweaty Palms: The Neglected Art of Being Interviewed, by H. Anthony Medley

Four books, reviewed by Ann L. Wiley and Helene Geiger 32

Communication and Education: Social and Psychological Interactions, by Garavel Salomon. Reviewed by Brockenbrough S. Allen 33

Designs for Instructional Designers, by Susan M. Markle, Reviewed by Wellesley R. Foshay 34


ERIC Reports on ID, Barbara B. Minor 36

About this issue...

JID is making the transition to a new editor and a new set of editorial and selection procedures. This issue contains articles chosen under the old and the new.

To make space for an abundance of interesting articles, recent issues of JID have slighted book reviews and ERIC reports on ID—two of our most popular features. We make up for it in this issue, which contains extra reviews and ERIC reports, covering many important works in the field.
A Comparison of the Leadership Behaviors of Instructional Designers in Higher Education and Industry

Barry Willis, Ed.D.
Director of Instructional Development
University of Alaska
Instructional Telecommunications Consortium
2221 E. Northern Lights Blvd. #210
Anchorage, Alaska 99504

Abstract. Instructional developers in industry and in higher education fill a number of leadership roles. Leadership is defined as the process of influencing the activities of an individual or group in an effort to accomplish situational goals. Though little research has been done, the literature suggests that the leadership behaviors of developers in higher education and in industry are quite different. Leadership behaviors of developers in these two settings were explored in terms of task and relationship behavior. The results indicate that differences in leadership may in fact exist, although not to the extent indicated in the literature.

Introduction

In years past, the majority of instructional developers were employed by institutions of higher education. This was due in part to the infusion of federal funds through programs such as the Fund for the Improvement of Post Secondary Education (FIPSE), which funded the creation of a number of instructional development centers in the early 1970s. In recent years, as federal funds appropriated for the improvement of instruction have all but stopped, the growth of instructional development efforts in higher education has slowed.

At the same time, declines in work productivity have encouraged industry to invest in employee training. While this training may benefit employees in enhancing their personal growth, it is most often designed to improve the technical competence of the individual. The task of developing this training has been assumed by instructional developers, or training developers as they are referred to in industry (Carriggan, 1980, p. 328).

In recent years, a number of studies have compared the competencies expected of developers functioning in these two settings. Although competencies of developers in higher education and in industry appear to be quite similar (Alexander & Yelon, 1972; Deden-Parker, 1981; Pinto & Walker, 1978), differences may exist in the way these individuals work with others.

Developer-client relationship building is frequently emphasized when instructional development is undertaken in an educational setting (Alexander & Yelon, 1972; Durzo, Diamond, & Doughty, 1979). In contrast, the highly competitive nature of industry often dictates that development efforts center on finding expeditious solutions to specific problems (Brancomb & Gilmore, 1975; Corrigan, 1980; Schwaller, 1980).

Whether the focus is on task accomplishment, relationship building, or a combination of both, developers in higher education and industry fill a leadership function, with leadership defined as the “process of influencing the activities of an individual or groups in efforts towards accomplishing goals in a given situation” (Hershey & Blanchard, 1972, p. 68). The leader’s facilitative responsibility, such as in the identification of group needs, the establishment of objectives, and the development of solution strategies, is a recurring theme (Bellows, 1959; Hershey & Blanchard, 1976; Koons & Donnel, 1968; Tannenbaum, Welsh, & Massarik, 1961).

Other leadership functions include “maximizing the utilization of group members’ abilities” (Schutz, 1963) and maintaining group cohesiveness (Stogdill, 1974).

Hershey and Blanchard (1977) suggest that in any leadership situation, a leader will display both task and relationship behavior, defined as follows:

**Task Behavior.** The extent to which leaders are likely to organize and define the roles of the members of their group to explain what activities each is to do and when, where, and how tasks are to be accomplished.

**Relationship Behavior.** The extent to which leaders are likely to maintain personal relationships between themselves and members of their group by providing socio-emotional support, psychological strokes, and facilitating behavior. (pp. 103–104)

The authors hypothesize that task and relationship behavior are separate and distinct dimensions of leader behavior and can be plotted on two separate axes as opposed to a single continuum, in which task and relationship behavior are looked upon as separate and distinct phenomena. From this perspective, an individual simultaneously exhibits high or low relationship behavior in a given leadership situation. As a result, four task/relationship behavior combinations are possible (see Figure 1).

The suggestion that situational factors play a key role in the leadership process is emphasized by Korman (1966), who concludes that task and relationship behavior alone have little predictive validity in identifying effective leader behavior. He states that validity might be improved through the concurrent exploration of task behavior, relationship behavior, and situational variables.

Extending the earlier work of Reddin (1967) and building on the work of Likert (1961, 1967), Hershey, Blanchard, and Hambleton (1977) have clustered situational variables under the heading of “effectiveness.” From this perspective, “effectiveness” is dependent on a leader’s ability to accurately assess the maturity of an individual or a group, with maturi-
ty defined as the "willingness and ability of a person to take responsibility for directing his or her own behavior" (p. 4). According to these authors, the maturity of an individual in relation to a specific job consists of two dimensions: Psychological maturity is related to the willingness or motivation to do something. Individuals who have high psychological maturity in a particular area of their work have the knowledge, ability, and experience to do tasks in that aspect of their job without the need for direction from others. (p. 4)

Job maturity is related to the ability or competence to do something. Individuals who have high job maturity in a particular area of their work have the knowledge, ability, and experience to do tasks in that aspect of their job without the need for direction from others. (p. 4)

This approach to the study of leadership behavior, labeled Situational Leadership Theory, suggests that the selection of an appropriate leadership behavior is dependent on the psychological maturity and the job maturity of an individual or group in a given situation. From this perspective, an individual or group possessing a great deal of psychological and job maturity requires a low task/low relationship behavior on the part of the leader. In contrast, an individual or group with little psychological and job maturity requires leadership behavior which is high on both relationship and task dimensions. Still other individuals/groups may have a high degree of psychological maturity, but little job maturity. Such an individual/group would require little personal nurturing, but a great deal of guidance in terms of the given task. As such, the developer's utilization of high task/low relationship leadership behavior with these individuals would be appropriate.

Despite the interrelationships which link leadership and the instructional development process, there is little known about the comparative leadership behavior of developers in higher education and industry. As a result, university-level instructional development programs are now training developers to function as leaders in higher education and industrial settings without a clear understanding of the leadership behavior of those currently working in these two settings.

This study was undertaken for the purpose of determining if there were any differences in developers' perceptions of their leadership behaviors. The specific leadership behaviors include task and relationship behaviors and leadership effectiveness as measured on a self-report leadership scale.

**Methods**

**Materials**

The Leadership Effectiveness and Adaptability Description (LEAD-Self) instrument developed at the Ohio State University Center for Leadership Studies (Hersey & Blanchard, 1977) was used to collect developers' perceptions of their leadership behaviors in different hypothetical situations. The LEAD-Self includes 12 hypothetical leadership situations. For each situation, four alternatives are presented. Respondents are asked to select the alternative which is most representative of their behavior in each situation.

Five scores are computed from the responses to the 12 leadership situations. Four of the scores reflect the number of times the respondent used each of the four task/relationship leadership behaviors (i.e., high task/low relationship, high task/high relationship, low task/high relationship, low task/low relationship). An individual's score on each of the four leadership behavior dimensions can range from 0 to 12. The fifth score reflects the respondent's situational "effectiveness" in selecting the leadership behavior strategy most appropriate to the psychological and job maturity of the group represented in each leadership situation. Scores on the "effectiveness" dimension can range from -24 (least effective) to +24 (most effective), ranging from -2 to +2 for each of the 12 situations.

**Procedures**

A stratified random sample was used to select 60 respondents from each of two sources: (1) instructional development division in higher education and (2) training departments in industry. A total sample of 120 was selected for the study. Seventy-five percent (N = 45) of the profiles sent to those in higher education and seventy percent (N = 42) sent to those in industry were completed and returned.

The means and standard deviations for the two groups in relation to the five leadership variables are presented in Table 1.

**Results**

Data were analyzed using one-way analysis of variance (ANOVA) with "setting" (i.e., higher education, industry) as the independent variable. Leadership "effectiveness"- and the four task/relationship leadership behavior combinations (i.e., high task/low relationship, high task/high relationship, low task/high relationship, low task/low relationship) were the dependent variables. The ANOVA was used to determine if there were significant differences among the five leadership behaviors reported by developers in industry and in higher education. In addition, omega square values (ω²) were computed for the F-ratio resulting from each of the analyses of variance.

The omega square value is a function of the significance of group differences and tells the researcher the proportion of variance accounted for by an individual variable or interaction. A small omega square value, for example, indicates that the difference in group means is small when compared to within-group variability. Although the variable may still be significant from a statistical perspective, its practical significance is minimal.

Developers in industry reported significantly more high task/high relationship behavior than developers in higher education, F (1, 82) = 3.91, p less than .05, although only 3.5% of the shared variance was accounted for by this variable (see Table 2).

Developers in higher education reported significantly more low task/high relationship behavior than developers in industry, F (1, 82) = 6.93, p less than .05, with 6.6% of the shared variance accounted for by this variable (see Table 3).

There was no significant difference in the mean number of times that high task/low relationship leadership behavior was reported in the 12 situations by developers in industry and higher education, F (1, 82) = 2.72.

The number of low task/low relationship behaviors used by developers in the
two settings was not significantly different, \( F(1,82) = 1.64 \).

As previously described, the respondent's "effectiveness" is directly related to the appropriateness of the selected leadership strategies to specific situational needs. The setting by "effectiveness" ANOVA indicates no significant difference between the leadership "effectiveness" scores of developers in the two settings, \( F(1,82) = .47 \).

**Discussion**

In terms of the 12 hypothetical leadership situations presented in the LEAD-Self, results indicate that developers in industry reported significantly more high task behaviors and significantly fewer low task behaviors than developers in higher education. The findings are in keeping with the literature suggesting that developers in educational settings emphasize relationship building as a major component of their dealings with clients (Alexander & Yelon, 1972; Durzo et al., 1979), while those in industry settings emphasize the need for task accomplishment and increased productivity (Branscomb & Gilmore, 1975; Schwaller, 1980). This difference in emphasis is summarized by Stolovitch (1981), who states that although the basic process of instructional development is quite similar, there are major differences between the worlds of the educational and industrial developer. Whereas the developer in education is concerned with learning and with individual growth of those involved in the development process, industry views development as a costly process and expects a significant return on investment either in terms of dollar savings or increased revenue.

Despite these statistically significant differences in task-oriented leader behavior, the relatively small omega square values reported in each ANOVA indicate that developers' reports of their leadership behavior in these two settings are quite similar. True developers in this sample from industry and higher education did have statistically significant differences on two of four task/relationship dimensions, but to say that these differences are extreme is to oversimplify the concept of leadership behavior in the two settings under study.

In terms of situational "effectiveness," there was not a significant difference in the leadership behaviors of developers in the two settings. This finding is not surprising in that the research literature did not indicate that developers in one setting are more effective than those in other settings.

As a result of this study, two conclusions may be drawn. One relates to the study of leadership behavior in general, the other specifically to the comparative nature of instructional development when conducted in higher education and industry.

First, although statistically significant on two of four task/relationship dimensions, a total of only 14.8% of the shared variance of developers in the two settings was accounted for. This suggests that either the leadership behavior of developers in industry and higher education is really quite similar, or that the data collection instrument used in this study is overly restrictive. The author contends that the two-continuum approach to the analysis of leadership behavior as used in this study is a significant improvement over the linear, single-continuum approaches often used in leader behavior research. Nevertheless, the results reported here suggest a need to further disembed task and relationship behavior in an effort to further define the individual components of these two dimensions.

In terms of the second conclusion, for a number of years researchers have been examining the competencies required of developers in higher education and industry (Alexander & Yelon, 1972; Dedon-Parker, 1981; Pinto & Walker, 1978) and comparing the two work environments (Corrigan, 1980; Schwaller, 1980). These studies, as well as others, indicate that although differences may exist, the functional gap between instructional development in these two settings may not be as great as once thought.

<table>
<thead>
<tr>
<th>Leadership Behaviors</th>
<th>Setting</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. High task/low relationship behavior*</td>
<td>Industry</td>
<td>1.32</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>Higher Ed.</td>
<td>.91</td>
<td>1.02</td>
</tr>
<tr>
<td>2. High task/high relationship behavior*</td>
<td>Industry</td>
<td>6.37</td>
<td>1.89</td>
</tr>
<tr>
<td></td>
<td>Higher Ed.</td>
<td>5.33</td>
<td>2.02</td>
</tr>
<tr>
<td>3. Low task/high relationship behavior*</td>
<td>Industry</td>
<td>3.80</td>
<td>1.96</td>
</tr>
<tr>
<td></td>
<td>Higher Ed.</td>
<td>5.00</td>
<td>2.18</td>
</tr>
<tr>
<td>4. Low task/low relationship behavior*</td>
<td>Industry</td>
<td>.51</td>
<td>.75</td>
</tr>
<tr>
<td></td>
<td>Higher Ed.</td>
<td>.74</td>
<td>.90</td>
</tr>
<tr>
<td>5. Situational &quot;effectiveness&quot;**</td>
<td>Industry</td>
<td>10.51</td>
<td>3.80</td>
</tr>
<tr>
<td></td>
<td>Higher Ed.</td>
<td>11.09</td>
<td>3.99</td>
</tr>
</tbody>
</table>

*Possible score of 12.

**Possible score of 24.

---

**Table 2**

Analysis of Variance Summary for Setting by High Task/High Relationship Behavior

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>22.71</td>
<td>1</td>
<td>22.71</td>
<td>.035</td>
</tr>
<tr>
<td>Within Groups</td>
<td>314.05</td>
<td>82</td>
<td>3.84</td>
<td>.05</td>
</tr>
</tbody>
</table>

*p < .05

**Table 3**

Analysis of Variance Summary for Setting by Low Task/High Relationship Behavior

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>29.09</td>
<td>1</td>
<td>29.09</td>
<td>.036</td>
</tr>
<tr>
<td>Within Groups</td>
<td>304.44</td>
<td>82</td>
<td>4.32</td>
<td>.05</td>
</tr>
</tbody>
</table>

*p < .05

Journal of Instructional Development
While we have developed a clearer picture of the specific competencies required of developers in these two settings, an understanding of the social/interpersonal component of the development process, of which leadership is but one important dimension, is more elusive. By gaining an increased awareness of leadership behavior and other social/task dimensions of the development process, instructional developers will be better prepared to make instructional development a viable and worthwhile activity, whether the context is industry or higher education.

References


Issues in Television-Centered Instruction for Adults

Penelope L. Richardson, Ph.D.
Assistant Professor and Director,
Center for Research on Adults
and Lifelong Learning, WPH 701
University of Southern California
Los Angeles, California 90007

Abstract. Interest in telecommunications delivery systems has increased in higher, postsecondary, and adult education institutions in recent years, in part due to demographic shifts and in part due to improved telecommunications systems. Current research on the adult learner and on instruction through media has grave flaws, and five reviews of research are needed to assist instructional developers and adopters to make wise decisions. A critical analysis is needed of existing telecourse packages. A review of research is needed on the motivation of various subgroups of adults and the recruitment and retention strategies which are effective for them. A review of research is needed on instructional strategies effective for different subgroups of adult learners. A review of research is needed on logistics and costs of alternative delivery technologies and audiences they are likely to reach. And a review of research is needed on the technical assistance and training needs of new adopters.

Introduction

Instructional developers today have exciting new options open to them as they develop their careers. Three societal trends are intersection, providing a time of opportunity for those professional enough and entrepreneurial enough to take advantage of it. The first trend is the boom of adult learners in our country today. Participation in college programs by adults beyond the traditional college age is expected to increase significantly, for a number of reasons. First, statistics show that people who have completed high school enroll more often in continuing education than people with less formal education, and the proportion of people over 25 who have graduated from high school is increasing, as is the proportion who have graduated from college. Second, as we move from being a goods-producing to a service-producing society, occupations are more likely to require more formal education. Third, since 1900 the work-life span has more than doubled. Men age 50 can expect to live to 75; women to 85. People can look forward to two twenty-year careers. Therefore many people in their fortieths and fifties are looking for educational opportunities that can prepare them for new jobs. Adult learners are the new market for well-designed instructional materials.

A second trend is the increasing willingness of institutions of higher education to collaborate in the development and delivery of learning opportunities to adults. Leaders in colleges and universities realize that technological developments in the workplace make it difficult for institutions of higher education to equip themselves with the most up-to-date facilities, and in some cases are therefore offering to do training in the workplace, taking advantage of the materials there. Continuing education agencies are collaborating with professional organizations to offer updated professional training in the variety of areas mandated by state legislation: nursing, dentistry, real estate, accountancy—the list goes on. Consortia of television stations and community and four-year colleges are collaborating to design and produce teletcourses, and in some states the equivalent of an Associate of Arts degree can be gotten through taking most of the courses "at a distance"—through multi-media packages of correspondence materials, television programs, audiotapes, and community-based seminars. These new collaborations of higher education institutions and community agencies mean that the adult learner has access to learning opportunities that are available at more convenient times and places and thus better able to fit into the working schedule of the busy adult.

A third trend is the marriage of television and interactive computers, so that interactive self-instruction is a viable option in the home and workplace. It is no longer necessary for the learner to transport himself or herself to the site of learning; instead, the information can be transmitted to the learner. We are only beginning to experience the ramifications of this development, but it has implications not just for formal higher education but for the learning of everything from home-plumbing to advanced statistics. Soon people will be able to purchase courses in a variety of topics at their local Radio Shack, and skip the higher education agency entirely, if they so desire.

What are the implications of all this for the field of instructional development? First, instructional developers need to look beyond campus-based instructional design and development centers as their primary source of employment. They should investigate materials development opportunities in various centers of enterprise: business and industry, professional associations, community agencies, health service agencies, and media centers. Adult learners are learning in a variety of places and for a variety of reasons, and formal higher education is only one of their options. Many will pursue their education from the workplace, the home, or the community center.

Second, instructional developers need to develop their broader human resource development skills. The jobs available in business and industry encompass more responsibilities than merely designing in-
struction. They include designing weekend seminars for managers, helping employees design career paths, developing instructional materials for customers. Jobs available in higher education for people developing multi-media materials require skills in networking, teamleading, political negotiations, and faculty development. Jobs in professional education require knowledge of that profession's value system, its communication patterns, its informal pecking order. Instructional designers who hope to be educational leaders need to develop themselves as managers, persuaders, and team-builders as well as excellent materials designers.

Third, instructional developers who plan to emphasize development of materials for formal higher education need to look beyond current practice and ask some critical questions. The telecourse development business has become something of a mini-industry, and developers have settled into a format adapted from the traditional college semester approach. Most telecourses consist of a certain number of television programs, an accompanying text and study guide, and a bank of test items. If telecourses are to continue to be developed, and in increased numbers, it is time for a critical review of the assumptions underlying the format currently used.

Increased clientele for adult learning, increased opportunities for institutional collaboration, and increased availability of new technologies of delivery—all promise a dynamic future for the instructional developer who broadens his or her human resource development skills, while asking the critical questions that will lead to improved practice.

In this country we have had decades of high hopes for the use of technology to solve educational problems, but so far, the printed word or picture is the most widely used technique for mediating instruction in higher education. Ten years ago, when there was great optimism about applications of electronically mediated instruction, the Carnegie Commission on Higher Education (1972, p. 1) forecast that “by the year 2000 it now appears that a significant proportion of instruction in higher education on campus may be carried on through informational technology.” But by 1975, a further study sponsored by the Carnegie Commission (Rockhart and Morton, 1975) suggested instead:

...the real impact of the new technology will for the most part be adding to, rather than replacing, current learning mechanisms. Some current mechanisms will be displaced, but the new technology will allow two major improvements. First, it will provide increased enrichment... Second, it will provide increased access by university faculty to students formerly outside the geographic limits of the regular educational system (p. 276).

It is this latter possibility that is now generating such enthusiasm on the part of certain higher educators and educational technologists. Demographics indicate that colleges and universities of the 1980's and 1990's will have to deal with an increasingly diverse range of students—diverse in age, educational purpose, background and preparation, socioeconomic status, and ethnicity. If those institutions are to meet the nation's educational needs in the future, so the argument goes, they cannot limit their concern to students who are between the ages of 18 and 25, white, middle-class, and academically skilled (Chickerling, 1981). Major pressures for change come from nontraditional groups of students new to higher education: those over 25 (who already outnumber students under 25), women with families, minority students, and students from non-academic backgrounds. Clearly, to the enthusiasts for technology, the problem can best be solved by telecommunication technologies, which can deliver instruction to students at times and places convenient to them.

Participation in college programs by adults beyond the traditional college age is expected to increase significantly, for a number of reasons.

A number of events nationally have heightened the sense that, at last, it's all about to happen. Many telecourse consortia are now in the business of producing, marketing, and offering telecourses (Munshi, 1980; Richardson, 1980; Yarrington, 1979). A study done in 1979 by the National Center for Education Statistics and the Corporation for Public Broadcasting indicates that some 71% of the nation's higher education institutions made some use of television for instructional and other purposes during 1978-79 (Dirr, et al., 1981). The Public Broadcasting Service began its new Adult Learning Programming Service in September, 1981, offering several telecourses each semester through public television stations to students unable to attend college on campus. New technologies are proliferating, and the 1978 Public Telecommunications Financing Act redefined telecommunication to include many modes of delivery: transmission via broadcast television and radio and other-than-broadcast distribution including coaxial cable, optical fiber, broadcast translators, satellites, discs, microwave, or laser transmission through the atmosphere. The Carnegie Corporation has funded the American Association of Higher Education to establish a Center for Learning and Telecommunications (Center for Learning and Telecommunications, 1981). And to clinch it and convince everyone that the Promised Land is finally here, former ambassador Walter H. Annenberg has given $150 million dollars to the Corporation for Public Broadcasting, to create new telecourses and to demonstrate the use of communications systems in solving higher education problems (Note 7).

Clearly, this proliferation of technologies and accompanying surge of entrepreneurial activities raises questions of great importance to higher and adult educators. What criteria should be used for selecting and supporting programs and projects? What is known about adult motivation and recruitment practices effective with different subgroups? What is known about instructional strategies effective in bringing about learning achievement for adults who participate in mediated instructional systems? What is known about logistics, costs, and audiences for various technological delivery systems?
What is known of the technical assistance and training needs of the new adopters? Decision-makers who turn to the research on mediated instruction or on adult learning will find little to guide their decisions. Wilkinson's critique of sixty years of research on media in instruction (1980) summarizes a number of problems with the research: imprecise definitions of terms, poor research questions, faulty experimental designs, research which trades off internal validity for external validity. Cross's attempt to make sense of research on adults as learners (1981) indicates a similar paucity of useful findings, probably because age per se is not a critical variable in predicting learning (Knox, 1980). Much of the research she summarizes was not done on learning outcomes at all, but on reasons adults gave for participating or on claims they make about their "self-directed learning projects."

4. A review of research should be done on the logistics and costs of alternative delivery technologies and support systems and the audiences they are likely to reach.

5. A review of research should be done on the technical assistance and training needs of the new adopters.

I. A Critical Analysis is Needed of Existing Telecourse Packages

As the certainty grows that the modern American college will be populated by a very different kind of student, the adult learner, much thought is being given to implications for the mission of higher education. Chickering (1981) has suggested that the idea of human development can supply a unifying purpose for the modern American college.

By understanding how students from 16 to 80 meet life cycle challenges and grow in terms of intellectual competence, ego development, moral and ethical development, humanitarian concern, interpersonal competence, capacity for intimacy, and professional development, educators can examine the potential contributions of various disciplines, areas of professional preparation, and educational practice. (Chickering, 1981, p. 7)

In the development of telecourse materials for adults, it is especially important that a clear definition of the criteria for excellence be developed. New technologies are developing rapidly, new monies are available for course development, and new pressures are being felt by technologists and educators to produce and market their wares. The various institutions involved in telecourse development are beset by a variety of political pressures and could easily be tempted by quick profit, pulled a dozen different ways and lose sight of the key question: the purpose of education.

Conceptually, the kinds of questions which must be asked about existing or developing telecourses include the following: What is their view of the learner? Of the teacher? Of the interaction between the two? Of the structure of knowledge, skills, and attitudes it is hoped the learner will develop? Of the relationship between the goals of the learning situation and broader societal needs? Course producers need to make explicit and to subject to public scrutiny the theories of content, teaching-learning process, and short- and long-term student outcomes (Putman and Retcher, 1973) upon which instructional choices are based. A constant dialogue between theory and practice is needed as the work of course development and acquisition proceeds.

Koppell and Chickering (1981) question whether educational technologies and varied forms of mediated instruction can contribute significantly to life cycle developmental tasks and to major dimensions of adult development. They concede that such vehicles may do a reasonably good job in imparting knowledge, communicating concepts, and helping students learn the language, methods of inquiry, and modes of thought of various disciplines or vocations. They recognize that such media may, when properly developed, achieve a wide range of training outcomes, especially if they can respond effectively to individual differences. But they view skeptically the suggestion that such vehicles could aid students in more general areas of human development.

A critical analysis of existing telecourse packages could identify those examples which seem aimed at accomplishing the broader adult development goals mentioned above, and could clarify and make explicit the different curricular conceptions represented in the various packages (Vaillancourt, Note 1). The packages could also be evaluated according to the degree to which they assist learners to identify their own strengths and weaknesses as independent learners, and provide developmental assistance, as suggested by Richardson (Note 2). The outcome would be the development of evaluative criteria to guide the creation of new course packages, criteria which concern the new goals for higher education and the importance of adult development as well as learning achievement outcomes.

II. A Review of Research Should Be Done On The Motivations of Various Subgroups of Adults and the Recruitment and Retention Strategies Which Are Effective for Them.

One of the hopes of funders of distance learning opportunities is that...
access will be provided to higher education for groups of students previously unable to take advantage of educational opportunities. Data on who currently participates in the distance learning opportunities available is scattered. Bryan and Forman (1977) indicate that the population of University of Mid-America distance learners is 75% women, average age 37, median income $11,000, half of whom had not participated in formal education during the previous five years. Their goals were personal satisfaction and career development, and two-thirds wanted degrees. McIntosh and Woodley (Note 3) indicate that Britain's Open University appeals mainly to the upwardly mobile students of working class backgrounds, who have professional and technical jobs. The labor class is unrepresented (true in America too), although the original purpose of the British Open University was to provide a "second chance" for the educationally deprived.

A review of research is needed to identify strategies effective in recruiting and retaining different subgroups of adult learners, so that decision-makers have more than demographic data to go on.

Such a review should have a number of characteristics. First, it should provide operational definitions of the strategy in question. Much that has been written about recruitment is vague and exhortatory; recruitment strategies such as "media advertisements," "alumni networks," or "college fairs" are mentioned. But from much that is written, it is not clear who the strategy is intended to reach, why people believe it would be successful, and specifically what the psychological functions of the recruitment strategy are. The same is true for retention strategies. Explicit operational definitions of recruitment and retention strategies need to be gathered.

Second, such a review should include the literature on effective recruitment and retention strategies for campus-based learners as well as distance learners, and for other kinds of activities besides learning activities. The basic question in recruitment and retention is this: What causes people to decide to participate in something and to stick with it once they're involved? In an Education Testing Service summary of several studies on college student attraction and retention, the reasons students give for leaving include academic matters; financial difficulties; motivational problems; personal considerations; full-time jobs; the expressed need for new, practical, non-academic experiences; and the lack of initial plans to obtain a degree (Ramist, 1981). Some of the reasons don't fit the adult distance learner; others do. Specifically, two factors crucial to retention were faculty-student interaction and effective counseling. Much has been written about campus-based recruitment and retention, and this literature should be culled for useful advice for planners of distance learning.

Third, such a review should provide a functional analysis of the recruitment/retention strategies presented. For example, the psychological literature tells us that the outcome of persistence is linked to the strategy of reinforcement. Whether on or off-campus students are the target audience, activities which reinforce for them their efforts are needed if they are to persist as students.

Fourth, such a review should identify the critical characteristics of adults which predict participation and persistence in a distance learning program. Cross (1981) has developed a Chain of Response Model which summarizes key factors others have suggested are significant. She suggests that when a learner needs to learn something, and realizes it, attitude. Also, the individual differences of adults should be taken into account in such a review. What attracts and retains a successful and self-confident learner might be very different from what would attract and keep an anxious and inexperienced learner.

Fifth, such a review should be theoretically oriented, showing the relationship of recruitment/retention strategies to psychological effects on the learner. Miller's force field analysis (1967), Rubenson's expectancy-valency paradigm (Note 4), Bosher's congruence model (1973), and Tough's anticipated benefits model, are all attempts to develop explanatory principles for adult participation. A theory of the achievement motivation and cognitive development of adults of various aptitudes needs to be developed and related to the effective environmental interventions (recruitment/retention strategies) which might be used.

Sixth, such a review should be eclectic and pragmatic, making appropriate use of strategies grounded in marketing, research, social psychology, and communications theory. Mass commu-

Since 1900 the work-life span has more than doubled. Men age 50 can expect to live to 78; women to 83. People can look forward to two twenty-year careers.

1983, VOL. 6, NO. 3

---

tions researchers and practitioners (such as advertisers) have a good deal of knowledge applicable to recruiting adult learners to telecommunications, and that knowledge should be tapped.

Lavish expenditures to develop new learning opportunities are wasted if adults don't take advantage of them, and "equal opportunity" arguments are empty if only the affluent and educated are customers for distance learning. A review of research on recruitment and retention strategies for different subgroups is needed.

III. A Review of Research is Needed on Instructional Strategies Which Predict Learning Achievement for Different Sub-Groups of Adult Learners.

One of the problems with the newer telecommunications technologies is that their glamour often prevents people
from thinking clearly about them. As Clark points out (Note 6), one result is that developers too rarely distinguish between a technology as a mode of transmission of messages, instructional or otherwise, and technology as a set of specific techniques employed in the instructional act. "Technologies of transmission" refers to those technologies which transmit messages. They are simply delivery devices which transmit instructional programs, in the same way that a postman or a carrier pigeon delivers a letter. "Technologies of instruction," on the other hand, are the set of research-derived principles for the design of instructional strategies which make up the teaching-learning process; strategies such as cueing the learner so that he/she attends to the appropriate instructional task, designing the instructional message so that it activates existing information processing skills, reinforcing and providing feedback, or building appropriate amounts of novelty or humor into instruction. Wilbur Schramm (1977) has recently noted that learning and achievement are more affected by the techniques we employ in instruction. Thus the findings are difficult to interpret. It has been commented that educational psychologists are able to measure individuals with micrometers, but that they tend to measure treatments with divining rods, and that situation still prevails today. A number of individuals are working on developing better specifications of instructional treatments (Leinhardt, 1980; Clark, Note 6) and their work can be consulted for further guidelines in this area.

Second, such a review should include the literature on effective instructional strategies for all age groups, not merely adults. Many research-derived principles about human learning and instruction have been discovered which offer guidance for the design of instructional strategies for all ages. Strategies such as cueing the learner so that he/she attends to the appropriate instructional task, designing the instructional message so that it activates existing information processing skills, reinforcing appropriate behaviors, or building appropriate amounts of novelty or humor into instruction (Menger, 1981). These findings need to be applied to instructional design for adults rather than putting adult instruction in a ghetto uninfluenced by discoveries about learning.

Third, such a review should provide a functional analysis of the instructional strategies presented. (Salomon 1979) describes three models for thinking about the function of all methods: remedial (in which the function is to ensure that the learner acquires the skills he/she needs in order to solve an instructional problem), compensatory (in which the function is to circumvent a learner's weaknesses by supplying him or her with the temporary aids needed in order to solve an instructional problem) and preferential (in which the function is to build upon the learner's special capabilities). Each of these models is supported by a body of instructional research which suggests helpful design principles to practitioners.

Fourth, such a review should identify the critical aptitudes in adults which predict learning from a given instructional strategy. Adults with high general ability, well-developed cognitive strategies, prior learning in the subject area, constructive (as opposed to defensive) motivation, and a strong sense of self-efficacy are likely to do well in an instructional situation which puts much of the information-processing load on them, in other words, one which is relatively unstructured; adults who lack these characteristics are likely to need their instruction to be more directive, structured, broken into small steps with immediate feedback at each step, and so on. Further writing on the relationship of adult aptitudes to appropriate instructional techniques can be found in Gronbach and Snow, 1977; Clark, Note 6; Allen, 1975; Goldstein, 1973; Salomon, 1979; Richardson, Note 2).

Fifth, such a review should be theoretically oriented, showing the relationship of instructional strategies to psychological effects on the learner. (Salomon 1979) points out that we need to move research on media in education from an atheoretical, unsystematic, "isolationist" exploration to a more systematic, theoretically guided inquiry, one which would relate the learner's information-processing capacity to the medium's particular way of structuring and conveying contents through its symbolic systems.

If media differ along specific dimensions (such as their technologies of transmission, contents, or symbol systems), then there are a number of possible ways in which media can interact with human behavior and learning. Not only do the sources of variance differ, but they affect different domains or outcomes. The technology of a medium affects the modes of interaction with users (compare computer-based instruction with a television show of the same content), and the transmitted content affects the knowledge acquired. The societal context in which a message is received affects emotional states, and the didactic mode of the message affects the ease of processing its content. It follows, then, that the different aspects of media interact with different aspects of behavior, thus undermining a conception of media as invariant entities. (Salomon, 1979, p. 8)

Salomon's Interaction of Media, Cognition and Learning argues that study is needed of the nature of a
medium's most essential attributes and the psychological functions that they can be made to accomplish under different conditions and for different learners.

Sixth, such a review should be eclectic and pragmatic, making appropriate use of strategies grounded in behavioral, cognitive, and humanist learning and instructional theories, as opposed to being based in a narrow philosophical viewpoint. It is likely that for certain learners, given certain goals and learning tasks, one set of instructional strategies might be quite appropriate, and for a different group, entirely different strategies would make more sense. Goldstein’s Structured Learning Therapy (1975) proposes a four-stage instructional sequence of modeling, role-playing, social reinforcement, and transfer of training to help mental patients learn community living skills. Such strategies are clearly grounded in behaviorist and social learning theories, and are appropriate for a client group with a restricted language code (Bernstein, 1964), an authoritarian upbringing, and a low sense of self-efficacy. For the motivated group of middle managers attending workshops given by adult educator Malcolm Knowles (1977), the humanistically-oriented techniques he proposes of collaborative goal-setting, planning, and evaluation make sense, since he's dealing with a group of individuals who already have the cognitive strategies to engage in self-directed learning. For instructional situations in which the goal is meaningful learning or concept acquisition, strategies such as cueing, defining, providing examples and nonexamples, sequencing, providing practice for mastery and for transfer, and reinforcing, are all appropriate (Merrill and Tennyson, 1977). A good review of instructional strategies should be practical rather than ideological.

Seventh, such a review should be especially clear about the counterintuitive findings contained in recent instructional research. For example, a favorite saw of adult educators is that adults learn more if you allow them to choose the instructional method; however Clark (Note 6) has reviewed a large number of studies which indicate that learners tend to select the method from which they will learn the least. Another "obvious" truth for a while was that adults who are slower learners will profit more from novel, media-oriented approaches. In fact, the research-based finding is that slower learners need more structure and sequence, while the capable students can handle more novel and less-structured approaches (Cronbach and Snow, 1977). Another incorrect assumption has been that if you want to predict who will learn most, it's important to know such things as sex, socioeconomic status, religion, and goals. In fact, general intelligence is the best predictor of amount of learning, and when that is controlled for, the sociodemographic factors don't account for any variance. The sociodemographic factors are useful in helping to predict who will participate, but not who will learn. Gage and Berliner (1975) provide an excellent beginning list of counterintuitive findings with implications for appropriate use of instructional strategies.

Such a review of research on instructional strategies would be a contribution to the field of adult instruction in general, not just to the field of television-centered instruction. Comparative media studies indicate that the technology of transmission is not a significant variable in learning outcomes (Schramm, 1977; Briggs, et al., 1967). Any medium can be used to teach, and the selection of strategy should rest on our current knowledge of instructional research and development, which would be summarized in a good review of research on instructional strategies.

Lavish expenditures to develop new learning opportunities are wasted if adults don't take advantage of them.


Such a review should have a number of characteristics. First, it should catalog the various technologies, give a functional description of each, characterize its critical delivery capacities and limitations, and provide an estimate of costs. Dordick et al. (1979) have developed a manual entitled, "TTV: A User's Guide" detailed case studies. For example, the TAGER Television Network links seventeen college and university campuses in Texas to a large number of corporate subscribers in the area, who receive graduate courses in business and engineering and handle around 2,000 registrations per semester. In Seattle, a consortium of eleven colleges and universities is providing credit and non-credit courses, professional continuing education and developmental education to 100,000 homes and public libraries by
cable (Dirr, Note 8). In Ohio, a consortium is using the two-way capabilities of the Warner Amex Qube system to provide courses (Dirr, Note 8). In San Diego, the State College is using cable radio to provide credit courses which fulfill the general education requirements leading toward a bachelor's degree (Dirr, Note 8). Many community colleges are collaborating with public broadcast stations and other community agencies to serve both formal and nonformal learning needs of the community (Richardson, 1980; Yarrington, 1979). These case studies should include empirical data on cost, time, personnel, and management systems involved in the various models, and should clearly identify the economies of scale which result from the different models. The case studies should relate the technologies to such features of their setting as community size, institutional factors, and alternative uses. They should also be explicit about unintended consequences.

V. A Review of Research Needs to be Done on the Technical Assistance and Training Needs of the New Adopters.

The Corporation for Public Broadcasting's Higher Education Utilization Study (Dirr et al., 1981) found that major factors which seem to affect positively the use of television for instruction include: faculty members' support for the use of television for instruction, availability of courses which meet the academic needs and standards of the colleges, and television outlet operators who are sympathetic to the colleges' goals for television use. Each of these areas imply a need for technical assistance and training.

Faculty members' support for the use of television for instruction can best be gotten by involving them in all aspects of decision-making. Faculty members' support for the use of television for instruction can best be gotten by involving them in all aspects of decision-making. Evaluation of the federal programs developed during the U.S. "Great Society" years of educational reform and of federal attempts to change the school indicates that the local practitioner is a powerful and legitimate force to be reckoned with (Dollar, 1978). Research on adoption and diffusion of innovation (Havelock and Huberman, 1978) argues for the importance of the local educator or planner. The complexities of the educational setting as a social system, combined with the primacy of personal contact, argue for the necessity for local faculty members to participate in all aspects of program development or utilization. Effective implementation requires personal interaction and contacts among practitioners and between practitioners if old roles are to be changed and new ones are to be learned.

Empirically, the overall quality of any instructional development process will ultimately be judged by the quality of the instructional products and the teaching-learning system developed. Instructional products must be developed which meet the criteria for quality in these three areas: content, instructional design, and aesthetic appeal. Additionally, they must meet the needs and expectations of external critics who act as gatekeepers. Richardson (1980a) found that different criteria are employed by people who hold different roles in an organization. Faculty members look for "materials which have substantial educational content and meet the standards of the college." Station managers are more interested in "materials which are of high production quality and meet the standards of the station." Instructional designers are most concerned that "materials are structured to enhance motivation, participation, persistence, and learning on the part of the at-home learner." Marketers are looking for "materials which must be designed as a coordinated instructional system, with all components working together to serve the same instructional objectives. Local support services must foster involvement (personal contact by phone or mail, seminar opportunities, counseling) and feedback (self-study quizzes, prescriptive comments on homework). Administrative arrangements must be designed for the adult learner: information and recruitment materials must be easily available; simple procedures must be available for registration, taking exams, obtaining course credit. Television programs must be offered at times convenient to the adult learner. These criteria were developed through interviews with telecourse students, faculty members, station managers, community college administrators, and marketing representatives, and need further validation (Richardson, 1980a).

All of the above topics have implications for the technical assistance and training needs of the new adopters of telecourses, whether they are designing, selecting, or establishing support services for them. The new technologies
retain students, had an inadequate grasp of the teaching-learning process and the related effective instructional strategies, and weren't able to train and develop their staff quickly enough to serve their students. So many of them withdrew away shortly. It wasn't enough to then deliver students to the instruction, and it isn't enough now to deliver the instruction to the student. We need to take stock before we charge ahead, and some reviews of research for decision-makers can assist us in the stock-taking.

Reference Notes

References
Bernstein, B. Social class, speech systems and psychotherapy. British Journal of Sociology, 1964, 15, 54-64.
Measuring the Importance of Learning from Instruction

The Outcome-Consequence Model of Evaluation

Michael J. Hannafin Assistant Professor University of Colorado Boulder, CO 80309

Abstract. Traditional models for the evaluation of instructional products and systems tend to be unnecessarily narrow. Often, the importance of learning beyond the immediate instructional setting is either not emphasized or completely overlooked. The Outcome-Consequence Model (OCM) emphasizes the importance of evaluating instructional products and systems using a broadened perspective of the intent of learning from instruction. Specifically, OCM emphasizes evaluation procedures that are systematic and empirical, establish legitimate and appropriate bases for comparison of instructional effectiveness, and evaluate the consequences of learning beyond the immediate temporal environment. While traditional evaluation models may be simpler, the consequences of learning are often not considered when using such models. Instructional designers and evaluators need to direct more attention to the importance of learning, and not simply the act of learning from instruction.

Evaluation of instructional products and systems has been the focal point of considerable study during the past 15 years (Alkin, Kosciof, Fitz-Gibbon, & Seligman, 1974; Davis, Alexander, & Yelon, 1974; Hively, Maxwell, Rabehl, Sension, & Lundin, 1973; Worthen, 1977). Generally, the instructional development profession appears to be in substantial agreement regarding the broad classifications of the types of evaluation, formative and summative, applied to typical instructional development projects (Dick 1977a, 1977b; Kaufman, 1972; Gagne & Briggs, 1979; Briggs & Wager, 1981). However, questions regarding the value of much traditional evaluation activity have been raised (Class, 1976; Guba, 1978; Hannafin, in press; House, 1976). The problem, it appears, is related not to the broad classification of evaluation, but to the specific conceptual validity of evaluation questions raised, the manner in which such questions are operationalized, and the limited scope with which instructional products and systems have been evaluated.

The purposes of this paper are:
(1) to expand the notion of evaluation to include both intuitive and empirical planning considerations;
(2) to examine the "tunnel vision" effects of simple outcome-based evaluation; and
(3) to propose an empirical evaluation model, the Outcome-Consequence Model, that includes not only the immediate outcomes of instruction, but the consequences of instruction as well.

The Definition of Evaluation Revisited
Evaluation has been defined and redefined continually throughout the past two decades (see, for example, Ebel, 1965; Gagne & Briggs, 1979; Popham, 1975; Provus, 1969; Stake, 1967; Stufflebeam, Foley, Gepphart, Guba, Hammond, Merriman, & Provus, 1971; Thorndike & Hagen, 1969). One of the least cluttered definitions of evaluation has been offered by Worthen and Sanders (1973), who define it simply as "...the determination of the worth of a thing" (p. 19). It is precisely the determination of the "worth" of instructional programs that has led to a broadened interpretation of the intent and purpose of evaluation. Worth, from an instructional standpoint, may be perceived as acquiring a skill, changing an attitude, or any of a variety of other learning outcomes. It has seldom, however, been interpreted within the context of the value, in either a qualitative or quantitative sense, of learning outcomes beyond an immediate instructional setting. It is for this reason that the Outcome-Consequence Model (OCM) of instructional product evaluation has been devised.

Outcome-Consequence Model (OCM) of Instructional Product Evaluation
The Outcome-Consequence Model (OCM) is a systematic procedure which is used to evaluate both current learning outcomes and the implications of these outcomes across time. Outcomes may be considered as the acquisition of an attitude, performance of a skill, or other learning outcomes of interest. The rationale underlying OCM is that instructional effects are best considered within a context that permits the evaluator to examine the consequences of learning, given past, present, and future influences. In order to examine such consequences, the relationships among antecedent, current, and future outcomes must be systematically identified and operationalized. The OCM evaluator:
(1) establishes a systematic and empirical approach to the examination of evaluation questions of interest;
(2) establishes a legitimate basis for performance comparison;
(3) considers the consequences of learning beyond the immediate instructional setting. The rationale, purpose, and characteristics of each OCM component are outlined in Table 1, and described below in detail.

Systematic and Empirical
Traditional evaluation is often so simplified that highly relevant questions regarding representativeness of groups, performance variance, and the effects of learner characteristics on product effectiveness are not considered. In OCM, the manner in which the intuitive


<table>
<thead>
<tr>
<th>Component</th>
<th>Rationale</th>
<th>Purpose</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0 Systematic and Empirical</td>
<td>Evaluation should be an orderly, systematic, and operationalized plan of both underlying assumptions and resulting products.</td>
<td>To provide a pre-formulated, empirical set of anticipated outcomes, including both performance-based and attitudinal considerations, in quasi-hypothesis form.</td>
<td>Anticipated outcomes—outcome-based attitudes and interactions are specified in a causal manner to locate more accurately the “real” sources of effects.</td>
</tr>
<tr>
<td>2.0 Basis for Comparison</td>
<td>Evaluation should reflect meaningful and appropriate comparisons among legitimate instructional objectives, not simply performance vs. non-performance.</td>
<td>To determine, for informed decision-making purposes, the actual increments in obtained performance compared with pre-existing or readily available alternatives.</td>
<td>Obtained outcomes are compared with assumed zero-level performance, existing knowledge levels (e.g., pretest, performance in previous instructional units, etc.), and expected levels of performance based on existing instruction.</td>
</tr>
<tr>
<td>3.0 Temporal Considerations</td>
<td>Consequences and meaning of current learning are best understood when the effects of prerequisite skills on current skills and current skills on subsequent, related skills are systematically evaluated.</td>
<td>To systematically evaluate the effects of different levels of supposed prerequisite skills on current performance, and to determine the carry-over effects of current learning on subsequent, related learning.</td>
<td>Reflects the reality of instructional installation, and projects the consequences of the existence, or non-existence, of prerequisite skills. Also, the extent to which current learning effects subsequent learning under similar and dissimilar instruction is examined.</td>
</tr>
<tr>
<td>(a) Past Influences</td>
<td>Past, related, learning events—moderating influence on current learning—merely specifying supposed prerequisite skills neither establishes the extent to which such skills are actually needed, nor reflects the reality of typical instructional product or system installation.</td>
<td>To evaluate empirically the extent to which current learning is likely to be moderated by the presence or absence of pre-specified prerequisite skills.</td>
<td>Prerequisite skills—their presence, degree of presence, or absence—are “blocked” by objectives in matrix form. Judgments regarding the extent to which instruction can be provided with varying degrees of prerequisite skills are made.</td>
</tr>
<tr>
<td>(b) Sustained Effects</td>
<td>Performance effects resulting from product/system installation should be sustained even when subsequent instruction is provided under traditional models.</td>
<td>To evaluate the extent to which performance carry-over from current instruction affects the learning of subsequent, related information presented under traditional instruction.</td>
<td>The consequences of implementing novel instruction within traditional instruction is examined. Also, projected sustained differences—a goal of instruction—are systematically verified.</td>
</tr>
<tr>
<td>(c) Cumulative Effects</td>
<td>Performance under several related units of sequenced instruction—all under a new instructional format—should accumulate compared with either no new instruction or partial new instruction.</td>
<td>To evaluate the extent to which desired performance accumulates as a function of the amount of sequenced instruction delivered under new vs. old formats.</td>
<td>Expands traditional notions of product or system verification to include hypothesized effects of systematic “growth” under new instruction.</td>
</tr>
</tbody>
</table>

Judgments of instructional evaluators are operationalized is critical. Using hypothesis-like statements, the OCM evaluator converts intuitive or assumed beliefs regarding the instruction and associated effects into direct statements of anticipated outcomes. For example, the evaluation of a computational mathematics program, designed for low achieving math students, might be stated as “Concept A attainment for low achieving students is equal to or is greater than average achieving students under existing instruction.” In effect, the logic that guided design decisions is evaluated as well as the actual products and procedures. Learning or failing to learn Concept A, in this example, is formalized not as a single entity, but as a function of the assumptions and design considerations affecting the instructional system.

Basis for Comparison
Designers of instruction interpret evaluation findings as indications of product effectiveness. Often, this is accomplished by assessing individual or group performance on instructional objectives, with the criterion for effectiveness expressed as some percentage of correct learner responses. This is a sufficiently straightforward process, and quite appropriate for completely novel ID projects. In many cases, however, the need for instructional development becomes apparent after formal needs assessment finds existing instruction to be inadequate, i.e., learning has occurred under existing instruction, but not to satisfactory levels. In such cases, two bases of comparison are appropriate:

1. the obtained versus desired outcomes as typically determined, and
2. the difference, positive or negative, between the outcomes produced by the new versus the existing instruction.

When desired outcome levels have been attained, it may be of interest, but of little critical value, to determine formally the performance difference between new and old instruction. How-

1983, VOL. 6, NO. 3
ever, when desired levels have not been attained, the comparative value of new-versus-old instruction becomes important. In the previously described example of the mathematics program evaluation, the Concept A attainment of low-achieving students may have reached the level of desired attainment. However, when compared with concept attainment under existing instruction, the newly developed systems product may be far inferior. Unless an appropriate basis for comparison is established before-the-fact, conclusions regarding product effectiveness must be drawn based on a simplified learning-by-objective paradigm or on the subjective judgments of decision makers.

**Temporal Considerations**

Traditional evaluation models tend to treat skills as mutually exclusive islands of knowledge, placing little emphasis on the way prior learning influences current learning, or current learning influences subsequent learning. This is, perhaps, the point of greatest disparity between traditional and OCM evaluation models. When using OCM evaluation, a designer will systematically investigate how different levels of prerequisite skills affect new skill acquisition, and provide a way to evaluate how well new skills are sustained, and how well acquired skills affect the learning of subsequent skills. Instructional evaluation is broadened to emphasize the practical implications and consequences of instruction in a more realistic temporal manner.

**Past influences.** Virtually all instructional designers recognize that existing levels of knowledge influence the learning of subsequent, related information. In many cases, this recognition leads designers to specify prerequisite skills. While this is likely the least expedient approach, it is seldom possible to limit instruction only to those learners who meet rigidly defined prerequisite skill requirements—especially when designing large projects. Typically, instructional evaluation plans do not account effectively for the different levels of background skills likely to be encountered in a widely used instructional program.

In OCM, by contrast, use of assumed prerequisite skills can provide a fairly sophisticated multi-dimensional perspective to the evaluation of instructional program effectiveness. OCM considers both the ideals of product use (i.e., mastery of prerequisite skills), as well as the realities of product use (i.e., different levels of prerequisite skills), and the way both cases influence learning.

In the design, development, and evaluation of a large-scale mathematics system, for example, it is often useful to determine the legitimacy of assumed prerequisite skills in moderating related concept acquisition. In the case of computation of simple interest, the presence or absence of various prerequisite skills will undoubtedly affect the learner’s acquisition of the skill. It will do so, however, in a differential manner. Whole number multiplication, for example, will likely be far more critical than whole number-times-decimals multiplication—a skill that may actually be mastered through well-designed instruction in computation of simple interest. In practice, the OCM evaluator assesses the consequences of the presence, or absence, of salient prerequisite skills on the attainment of criterion information. As a result, the designer obtains realistic knowledge of the effects of different levels of prerequisite skills on criterion learning. Thus, it becomes possible to forecast the relative consequences of providing new

quotation and retention of the “missing skill,” but also with the transfer of that skill within the larger, intact system.

It is possible to design effective within-skill instruction that is structurally and procedurally incompatible with the existing program. For this reason, evaluation should include an indication of the cross-skill compatibility and transfer among instructional units. OCM emphasizes sustained effects in order to evaluate the extent to which initial learning has been subsequently retained, and the extent to which initial learning under the new system has affected subsequent learning under existing instructional systems of different philosophical, structural, and procedural formats.

**Cumulative effects.** Existing evaluation designs often include fairly detailed cross-tabulations of the performance of learners by objective, but seldom include a growth component. In OCM, the intent of instruction is evaluated according to anticipated cumulative effects as well as the performance-by-objective analysis used in most traditional evaluation plans. Mathematics learning, for ex-

Instructional designers must make certain that someone within the development team assumes responsibility for consideration of the meaning of instructional effectiveness beyond the immediate instructional setting.

---

Journal of Instructional Development
Instructional development. During recent years, however, a predictable territorial debate over important ID issues has often surfaced among instructional designers, traditional curriculum planners, media producers, and evaluators. Curriculum specialists, for example, have asserted that many of the temporal considerations advocated in OCM fall in the domain of curriculum planning, and not instructional design. Similarly, evaluators often perceive these considerations as evaluation-specific.

Ideally, ID projects should include appropriately placed and qualified personnel to permit role specialization. Under circumstances where such personnel are available, and where instructional development is managed effectively, and where important relational hierarchies have been identified, and where appropriate data exists regarding such skill hierarchies, and where curriculum and ID personnel work effectively as a team, the partitioning of responsibilities seems desirable. In many cases, however, ideal conditions simply do not exist. Therefore, instructional designers must make certain that someone within the development team assumes responsibility for consideration of the meaning of instructional effectiveness beyond the immediate instructional setting.

**Influencing Instructional Design via OCM**

It has been suggested that evaluation concerns may be more valuable during the formulation phase of an instructional development project than during the evaluation phase (Morris & Fitz-Gibbon, 1979; Hannafin, in press). It makes little sense to develop a sophisticated evaluation model to examine issues of interest when such issues were not considered initially during the formulation and production phases of development—effects are not likely to be found. On the other hand, careful consideration of the consequences of instruction and learning, applied during the early stages of development, is likely to provide instructional designers with additional relevant product specifications. When OCM considerations are made before-the-fact, instructional designers are more likely to be aware of issues such as across-skill instructional compatibility. In this regard, OCM is valuable in that an expanded formalization of instructional intent is articulated before-the-fact. Table 2 contains a series of questions typically applied during both the development and evaluation phases of each component of OCM.

While valuable during formulation and evaluation, such considerations seem most fruitful when an instructional program is first conceptualized and designed. Certainly, if some things are not considered early in the design of a project, many of the sources of data necessary for answering OCM questions will be either unavailable or difficult to collect. It seems at least as reasonable to use OCM as a guide to instructional planning as to use OCM as a procedure for instructional research.

**Closing Comments**

It should be apparent to the reader that OCM, as a procedure, will not apply in its entirety to all instructional system and product evaluation activities. However, it should also be apparent that the rationale underlying OCM is generalizable to virtually all instructional development and evaluation. Broadened planning, to include systematic consideration of the outcomes and consequences of instructional systems and products, is likely to improve the quality and meaningfulness of both development and evaluation.

OCM is not a new classification of evaluation; it is compatible with several existing evaluation models. Instead, the rationale and emphasis of OCM provide a broadened perspective on the process and purpose of instructional evaluation. The unique emphasis of OCM, however, requires instructional systems designers, developers, and evaluators to adopt an expanded perspective on the basic reasons for developing instructional programs.

For years, educators— instructional designers in particular—have been criticized for encouraging "tunnel vision" learning through the extensive use of objectives in materials design, implementation, assessment of student learning, and evaluation of the effectiveness of instructional products (Bisner, 1971; Meyen, 1970; Macdonald-Ross, 1973). On one hand, instructional developers argue that objectives, as such, do not necessarily foster restrictive learning (Popham, 1971). On the other hand, developers make distressingly few attempts to systematically disprove the criticism. The "big picture" of instruction—intent, effectiveness, consequence, influence—must be reconciled. As a profession, developers must not disregard the broader issues of instructional development. While the OCM model of evaluation is a modest step toward this awareness, the implications for changing the instructional environment are significant.

---

**Sample Considerations in Planning for Each Component of OCM**

1. **Systematic and Empirical**

   What are the specific hypotheses for outcomes of each portion of instructional product or system? What are the projected relationships among outcomes? What data will be needed to address each hypothesis? What are the assumptions underlying the design of instruction? What type of analysis, if any, will account for measurement errors?

2. **Basis for Comparison**

   What is the performance history of learners in the area of instruction? Will an overt be administered? What are the specific objectives for which instruction has been designed? What legitimate instructional alternatives exist? What instruction, if any, has been used to date?

3. **Temporal Considerations**

   (a) **Past Influences**

      What are the assumed prerequisite skills? What is the effect of the absence or presence of such skills on current skill acquisition? Which assumed prerequisite skills are most critical to the acquisition of current content? Can critical prerequisite skills be embedded into new instruction?

   (b) **Sustained Effects**

      For which subsequent instruction will current skills be most highly dependent? To what extent will the acquisition of skills via new instruction provide a continued advantage during subsequent learning? Does the variation in instructional format from new to existing instruction cause interference during the switching? Will transitional instruction, designed to blend new and old formats be helpful or necessary?

   (c) **Cumulative Effects**

      How many sequential instructional products/units/modules are to be developed? Which learners will participate in different phases of instruction? During tryouts, can learners be assigned to different combinations of new instruction, ranging from no new instruction to all new instruction? Will learning of prescribed content be affected positively by the amount of new versus old instruction?
toward this end, it represents a recognition of the important consequences of systematic instructional development.

References


Stake, R. The countenance of educational evaluation. Teachers College Record, 1967, 68, 523-540.


The only magazine devoted exclusively to instructional development, the Journal of Instructional Development (JID) is for those involved with the design, implementation and evaluation of courses or curriculum. JID contains articles on theories, techniques, reports, case studies, and critical reviews of instructional development projects and systems. It's designed to stimulate communication among instructional developers at all levels of education and training.
Notetaking Research

Implications for the Classroom

Carol A. Carrier
University of Minnesota
120 Peik Hall
Minneapolis, Minnesota 55455

Abstract. The study of student notetaking behaviors has produced useful insights into how students learn from lectures. This article presents five preliminary conclusions about notetaking practices based on findings from the notetaking literature. Each conclusion is followed by a discussion of the implications for classroom instruction. Finally, the author proposes links between various lecturer and student behaviors and the external events of instruction described by Gagne and Briggs (1979).

Notetaking: From Research to the College Classroom

Educational researchers have been accused of trying to prove through experimentation what everybody else has known for a long time. Researchers are most vulnerable to this accusation when their investigative targets are widely accepted practices or beliefs. Everyone can point to certain educational practices which are so familiar and commonplace that they seem automatic. When student or teacher practices become habitual, their effectiveness is rarely challenged. Instructional designers who question such practices on the basis of research findings learn that resistance to change is strong.

In recent years, educational researchers have begun to study one such widely-practiced behavior in the college classroom—student notetaking. Perhaps no study strategy would be more staunchly defended by students and teachers alike than that of recording notes while listening to lectures. Asking students to surrender their notebooks and pens at the beginning of a lecture is likely to incite a minor uprising. Instructors too would be uncomfortable. Most have grown accustomed to viewing a roomful of students busily recording information as a sign that students are actively engaged in learning from the lecture.

From a purely technological perspective, notetaking seems archaic. Inexpensive duplicating procedures make it feasible for instructors to provide students with copies of detailed lecture notes. Students can record a lecture on audiotape and listen to it again and again. But despite these conveniences, notetaking continues to be practiced in most college lecture courses. In a recent survey of one undergraduate sample, Carrier and Newell (Note 1) found that 100% of the respondents strongly agreed with the statement “I almost always take notes in lectures.”

How and why do students take notes? Does notetaking contribute to performance? In what ways can lecturers improve the environment for notetaking? Although many dimensions of notetaking remain to be investigated in more depth, sufficient research has been reported to allow us to extract some preliminary conclusions. These can serve as guidelines for instructional designers and other instructional improvement specialists who can help their faculty clients better understand the role of notetaking in college instruction. These faculty can in turn begin to explore how they might structure the classroom environment to help students take and use notes more effectively.

In the following sections of the article, five preliminary conclusions which emerge from the notetaking literature to date are presented. After stating the conclusion, a brief summary of research which supports it and the implications of the conclusion for college instruction are presented. Finally, suggestions are presented for how lecturer behaviors can positively affect notetaking practices within the context of the “events of instruction” proposed by Gagne and Briggs (1979).

Conclusion 1: Students who take notes during a lecture will learn more than those who simply listen.

The foremost question addressed by researchers in the early notetaking literature was to what extent does taking notes during lectures lead to better test performance than simply listening? The guiding hypothesis was that notetakers would learn more than non-notetakers because the former are more actively engaged in manipulating incoming information. Ladas (1980) provides greater clarity to this hypothesis by discussing notetaking in terms of current information processing models. These models propose a set of phases which characteristically take place in the learning of new information.

Initially, the learner must be motivated or aroused to selectively attend to the to-be-learned stimuli. Specific and general behaviors of lecturers designed to focus student attention and increase motivation illustrate this “orienting stimuli” phase of information processing. The lecturer might move to the podium and assume a particular posture which indicates to students that he is ready to begin. Or he might pose a provocative question, or tell a humorous story to lead into the lecture. If the lecturer succeeds in arousing student attention and interest, students will be more likely to display appropriate “orienting responses” such as putting aside the newspaper, taking out a notebook and preparing themselves to record information.

As the lecture is delivered, the serious
A. The Encoding Hypothesis

The audience is divided between Notetakers and Listeners (who do not take notes).

**Notetakers:**
- Attend carefully to avoid missing critical points
- Compare new information to stored information
- Translate lecturer’s words into own words
- Create a larger structure for later use
- Listeners, who do not take notes

**Daydream**
- Read other material
- Doodle
- Lose concentration

B. The External Storage Hypothesis

According to this hypothesis, the real value of notetaking occurs during later review of the notes.

**Typical Activities During Notetaking**
- Record in rote, verbatim fashion
- Copy, but do not react to or process information
- Record as much as possible without discriminating between essential and nonessential information

**Typical Activities During Later Review**
- Review and rehearse information
- Rewrite notes in own words
- Seek clarification of some points
- Compare prior knowledge with new knowledge

**Figure 1. Two hypotheses about notetaking**

### Conclusion 2: In general, students will learn more from a lecture if they both record and review their own notes.

As noted above, one hypothesis about why notetaking is more helpful than simply listening is that taking notes actively engages the learner’s attentional and processing functions during the period of listening. In other words, recording what the lecturer presents helps the student act upon new information at a deeper level. This viewpoint, then, stresses the encoding function of notetaking.

In contrast, others have hypothesized that notetaking is of value primarily because it results in a permanent record of ideas from the lecture and that this quality is more critical than possible encoding activities which take place during the recording of notes. Processing of any depth during notetaking is minimal; it is argued, because most students record information in a rote, verbatim fashion, failing to translate it into their own words. From this perspective, the real processing of lecture content occurs not during notetaking, but in later review.

To test which model best explains the effects of notes, studies have compared conditions which allow students to take their own notes with ones in which students are either given lecturer’s notes prior to or after the lecture. If a major benefit of notetaking is to facilitate active encoding of information, taking notes for oneself should be more facilitative than simply reviewing notes compiled by someone else. However, if the

### Implications for the Classroom

Studies contrasting listening with notetaking have consistently yielded support for notetaking practices. Some critics have claimed, however, that this research is not ecologically valid because conditions in the studies may be too dissimilar from those in actual lecture courses. To illustrate, some studies have used lectures whose content is unrelated to the course in which the study was run. Artificial constraints to not take notes or to listen only have been imposed on students who will normally take notes. The brevity of the lecture delivered in a study, sometimes lasting less than twenty minutes, is in contrast with the typical college classroom lecture which may run 45 to 50 minutes in length. But while any given study may indeed be weak in one or more of the areas mentioned, taken as a whole, this research represents the only systematic effort to empirically demonstrate the effects of notetaking.

Should college instructors encourage notetaking among their students? On the basis of the research to date, the answer to this question appears to be a definitive "Yes!" Further investigations of notetaking under natural classroom conditions will continue to enrich our understanding of this study strategy. Examining student notetaking behaviors across all lectures given within a single course would indicate if notetaking patterns differ at the beginning, middle, and end of a course. The effects of training students to use specific notetaking formats needs further exploration; some studies have noted a decrement in performance due to notetaking training (Corey, 1955; Palmatier & McNinch, 1972).
utility of notes rests mainly in their external storage capacity, externally provided notes should be as useful as personally encoded notes if the information contained within each is equivalent.

A study by Fisher and Harris (1973) is representative of how this question has been examined. They employed five treatment groups, each of which combined a notetaking strategy (i.e., took notes versus did not take notes) with a review strategy (i.e., reviewed own notes versus reviewed lecturer's notes versus mental review). While review of some type was useful under both notetaking conditions, subjects who were allowed to both generate and review their own notes scored higher on immediate free-recall and short term objective tests than any other group. This finding suggests that both the practice of taking notes and being allowed to review one's own notes are optimal study strategies.

Thomas (1978) offered several nonnotetaking groups a lecture summary to review while others reviewed only their own notes or their own notes in addition to the summary. Even though the lecture summary group members had more information units available at review time than the average member of a notetaking group, they recalled less. This study suggests that while notetakers may not have recorded the quantity of lecture content, some aspect of the coding process, the rehearsal of personally coded material, or both helped students retrieve information at the time of recall.

Implications for the Classroom

Clearly, the acts of taking notes and reviewing notes make individual contributions to the learning process. When students use idiosyncratic coding strategies, shortcuts or mnemonics as they take notes, reviewing their own notes may cue certain types of information or activate retrieval strategies. An instructor might encourage students to become more efficient in their notetaking by explicitly teaching some of these shorthand coding devices. Carrier and Newell (Note 2) found that less than 2% of the undergraduate sample they surveyed had received any instruction on how to take notes and more than half indicated they would like to take a course to improve their skills. It is possible that a small expenditure of time devoted to building notetaking skills could result in big gains in the quality of student notes.

Conclusion 3: Review of notes will lead to improved performance.

Review of notes was discussed above in the context of the encoding versus external store hypotheses. In this section, different types of review activities and their contributions to learning from lectures will be addressed.

The nature of review of notes is determined in part by the motivations of the individual notetaker. When it occurs, as well as the purposes for notes review, varies from student to student. Some students regularly review their notes following each lecture. In doing so, they intend to fill in gaps while the lecture content is fresh in memory, or to elaborate upon hastily noted words or phrases. Points which need further clarification from the lecturer are identified. Other students rewrite their notes after each lecture simply to make them "neater." Three-fourths of the sample in the Carrier and Newell survey, for example, indicated that following a lecture, they routinely add things to their notes and generate questions which they attempt to answer. The effects of specific strategies used during notes review is uncertain. In a study by Norton (1981), rewriting and rereading of notes seemed equally beneficial in promoting test performance, and both were superior to not reviewing notes at all.

This finding suggests that both the practice of taking notes and being allowed to review one's own notes are optimal study strategies.

A second form of review occurs immediately preceding a quiz, a midterm or final examination. This massed review grows out of the specific motivation to perform well on the impending exam. This activity is often referred to as "cramming" because the student hopes to retrieve as much as possible in a brief period of time for the sole purpose of examination performance. Here the student attempts to reconstruct an abbreviated version of the lecture to recall the lecturer's perspective on the content and to remember as many critical facts, concepts, principles, and the like. Most the delayed tests. These authors suggest that the benefits of notes/review can be attributed to its suppression of recall of irrelevant material, its potency to cue relevant information and the fact that it serves as an additional learning experience by providing more exposure to important lecture information. Only one study (Fisher & Harris, 1973) found that a review period before a delayed test was not facilitative. However, in this study, the review period immediately followed the lecture rather than being closely tied with the delayed test which was given three weeks later.
Implications for the Classroom

Instructors are reluctant to encourage cramming behaviors for examinations, perhaps because they fear that students will spend no time thinking about course content until a few hours before the exam. Although their fears may be warranted, massed review of notes before an exam has been shown to produce better recall and recognition of lecture content. These results are found regardless of whether the review period and the test occur a few hours or a few weeks after the lecture. Instructors can take comfort that if students review only before the exam, they will likely perform better, than if they did not review at all.

How massed review of notes compares to distributed review cannot be determined from existing research. If findings from other comparisons of massed and distributed practice generalize to notetaking review, we could expect greater gain for review when it occurs consistently following each lecture (Travers, 1997). To ensure that distributed review takes place, an instructor might build in a 5 or 10 minute review period at the end of each lecture or at the beginning of the next lecture. Norton (1981) found that despite good intentions, students may not carry through with regular review on their own. In her study, more than 90% of the students indicated on a prelecture questionnaire that they intend to revise their notes in preparation for exams. When queried after the examination given in the study, however, only 14% reported that they actually did revise their notes before the exam.

Conclusion 4: Lectures can facilitate student notetaking by highlighting important information and providing a clear organizational framework.

The quality of student notes and their subsequent worth to the notetaker during review will depend to some extent on the behaviors of the instructor. Students complain that it is difficult to take notes from some instructors, perhaps because they wander from one topic to another, fail to give cues, or speak so rapidly that no one can keep up.

As he listens to a lecture, the notetaker continually must decide what to record. Such decisions are influenced by many factors. Specific behaviors of the lecturer play an important role in this process. It has been found, for example, that items which are written on the blackboard have a high probability of being recorded in student notes (Maddox & Hoole, 1975; Howe & Godfrey, 1977). On the other hand, the use of some forms of visual aids actually reduces notetaking behaviors (Harley & Cameron, 1967). In part this can be explained because conditions of the lecture situation are changed (e.g., a darkened lecture hall during the presentation of a slide tape).

The pacing and organization of the lecture are also crucial. Even a skilled notetaker is capable of recording only a small proportion of what the lecturer presents. Ladas (1980) discusses a study by Greene (1928) who found that students took notes at the rate of 20 words per minute during a lecture. If a lecture is presented at the slow rate of 100 words per minute, the average notetaker will be able to record only about 20% of what is presented. What students do record, they tend to remember. There is a strong correlation between recording a piece of information in one’s notes and getting items relevant to that information correct on the test (Howe, 1970).

Does the organizational structure of the lecture affect notetaking? Logically, a salient lecture organization should be useful because the notetaker has to work less hard to recognize connections among ideas and to differentiate between essential and nonessential information. Research on organization of prose suggests that the importance of organization increases as the difficulty or unfamiliarity of the information increases. Scrambled passages have been shown to reduce recall over their logically ordered counterparts, but usually only if the to-be-learned material is difficult. Apparently most learners are capable of generating organizational frameworks where none are explicit.

Most of the studies investigating organizational variables and notetaking have had students read prose rather than listen to orally presented lectures. DiVesta and Gray (1973) manipulated the organization of prose by varying the degree to which six passage segments were related to a common conceptual superordinate theme and were contiguous with one another. They contended that the thematic relations among these passage segments would affect a subject’s selection of strategies for storage and retrieval of the information. The authors reasoned that notetaking should be more helpful in cases where there was little inherent organization in the material (e.g., six “discontinuous, unrelated” passages) than in materials with a salient organizational scheme (e.g., six “continuous, related” passages). They did not find the predicted interactions between notetaking and thematic organization, but found that notetaking led to superior recall under all organizational patterns.

Shimmerlik and Nolan (1976) hypothesized that encouraging subjects to reorganize passage material through notetaking would increase recall since it would strengthen associations or broaden access to that information at retrieval time. This concurs with predictions from the encoding variability hypothesis (Craik & Tulving, 1975) that multiple encoding strategies for the same information will result in greater remembering. Half of the subjects in the study heard the lecture presented using society as the organizing framework while the other half heard it with topics as the organizing framework. Further, half the subjects within each of the two organizations were asked to record their notes following the same sequential order as the passage while the other half were told to take notes using the alternative conceptual framework. In the sequential condition, then, those who heard the society organization took notes by society. In the reorganization treatment, subjects who received the
society organization were asked to order their notes by topic. The authors found that notetaking which involved reorganization led to greater free recall of information on an immediate test (Experiment 1) and on a delayed retention measure given seven days later (Experiment 2). They conclude that the utilization of a second organizational framework increased availability of cues during retrieval time.

Implications for the Classroom

The instructor who is intent on "covering all the material" in his lecture notes, regardless of how quickly he must speak, will find little encouragement in the notetaking literature. Notetaking is not an efficient process; students are able to record only modest amounts of information even at a comfortable rate of presentation.

An instructor can reduce the number of decisions a notetaker must make and thus increase efficiency by calling attention to important material. Verbally directing students to "remember this point" or "be sure to note the following" are useful highlighting strategies. The judicious use of blackboard and overhead transparencies is also recommended. Because students record most things written on the board, an instructor should use this medium strategically, taking care to avoid writing trivial or nonessential information. Making the organization of the lecture salient should also ease demands on students. The instructor should not hesitate to provide an overview, e.g., "In this session, I'll discuss five factors that contributed to the current economic recession in Brazil. For each, I will provide an example of the ways it operated at both the local and national levels." This strategy helps the notetaker establish a framework for high importance information, a strategy which is especially useful for certain types of students as I will discuss in Conclusion 5.

Conclusion 5: Students with different abilities and levels of prior knowledge may require different notetaking strategies.

Teachers would agree that students differ in their ability to participate in class discussions, to write term papers or to perform well on examinations. Similarly, different students have different notetaking styles and capabilities. Ganske (1981), after reviewing more than thirty notetaking studies, identified two major categories of notetakers. The first he labels "processors." These students tend to rephrase lecture points, doodle, miss key lecture points if no cues are provided, and often include irrelevant material. In general, the notes of "processors" are often incomplete and would be difficult for someone other than the notetaker to use. The second type of notetaker Ganske labels "transcribers." They characteristically record information verbatim and almost always note the major points in the lecture regardless of whether they have been cued. These notetakers rarely doodle and tend to be more wordy than processors. As a rule, their notes are more complete and have the appearance of being intended for future reference.

Although we can generate classification schemes for the various types of notetakers as Ganske has done, there is little empirical evidence to indicate that one style or approach is superior to another. Judging the quality of notes is complex business because how the notetaker uses his or her notes is unobservable and yet may be as important as the quality of the notes themselves. To an external evaluator, a set of notes may appear as a garbled collection of meaningless abbreviations and sentence fragments, but to the notetaker the notes may represent a web of links or associations to other knowledge or cues which will aid recall at a later time.

Rewriting and rereading of notes seemed equally beneficial in promoting test performance, and both were superior to not reviewing notes at all.

There is some evidence that certain students may actually learn less if they record notes during a lecture. Berliner (1971, 1972) examined how memory abilities relate to notetaking. He found that notetaking interfered with ability to recall information for students with poor short-term memory abilities. Di Vesta and Gray (1972) also found positive correlations between memory ability and performance on both immediate and delayed tests for notetakers. Peters (1972) found that notetaking during a lecture interfered with recall test performance for students with "low efficiency listening" scores as determined by their ability to recall words from a list presented at slow and fast rates.

In a study involving high and low math ability students, Peper and Mayer (1978) found that students who took notes remembered more interpretive-like or "far transfer" type of information while non-notetakers remembered more "near transfer" or generative items. These relationships were especially strong for lower ability subjects, causing the authors to suggest that lower ability students are less capable of assimilating new and unfamiliar information through listening. In a recently completed study, Carrier, Fligson, Klimoski and Peterson (Note 2) found that students with high levels of debilitating achievement anxiety took notes of lower quality than students with low debilitating anxiety or those with facilitative anxiety.

That students who are deficient in memory ability may require a different strategy than those with greater memory ability is consistent with other findings from research on individual differences and instructional manipulations. That is, learners with relatively low general ability, those with little prior knowledge of the instructional content to be taught or those with specific cognitive styles (e.g., field independence vs. field dependence) benefit from instructional techniques or methods which reduce processing burdens (Cronbach & Snow, 1977). In other words, to the extent that the instruction reduces the demand for student generated mediators, it will enhance the probability of success for these learners. An example of programmatic research supporting this conclusion can be seen in the work of Tobias (1976, 1981, 1982). Consistently, he has found that students with little prior knowledge of the content to be learned need high levels of instructional support. He defines instructional support as those conditions which, for example, enhance the organizational salience of the material, require overt responses or pro-
vide explicit feedback about responses. With respect to notetaking, giving explicit instructions to take notes, providing specific cues about what is important, and providing concrete suggestions on the quality of notes could be considered instructional support. Lecturer's notes are another type of instructional support which may be more beneficial than self-generated notes for some students.

Implications for the Classroom

Within most lecture classes, some students may be in need of extraordinary assistance from the instructor if they are to obtain a set of notes which can aid their studies. The research reviewed here suggests that some students will benefit from not taking notes at all!

The problem for the instructor is twofold: to identify, through feasible means, students who need special help, and to determine the most effective types of instructional support or mediating devices. Of course, an alternative to accommodating certain students is to provide everyone with these mediators. For example, detailed lecture notes could be supplied for everyone with the expectation that students who benefit from taking their own notes will continue to do so. Another mediating strategy, that of highlighting critical information, will be especially useful to certain students, but should not interfere with the processing behaviors of students who don't need it.

If an instructor is interested in accommodating the needs of certain subgroups, what steps might he take to do so? One initial action might be to collect and review the students' notes from several lectures. This would help him identify those students who seem to have inadequate notetaking styles. As a next step, he might hold individual conferences with these students to gain a clearer picture of explanations for their notetaking deficiencies. If he finds students with similar problems and needs, he might then plan interventions which would be appropriate. For example, if five students indicate that they find it impossible to follow the lecture because it moves too rapidly, they might be encouraged to listen to an audiotaped version of the lecture. If he finds that a number of students have difficulty discriminating between important and unimportant content, the instructor might provide some practice opportunities with feedback. For students who aren't aware of or skilled at using various shorthand techniques, he might demonstrate how these techniques can increase efficiency. These relatively minor interventions could lead to high payoff for some students in terms of improved notetaking and test performance.

Notetaking and the Events of Instruction

Gagné and Briggs' (1979) scheme for describing crucial "events of instruction" provides a useful framework for considering how learner behaviors can aid student processing activities. Table 1 displays this framework. Both research and practice can benefit from conceptualizing the external events of the lecture environment and the internal events which are student responses to this environment.

From an instructional research perspective, notetaking offers an important vehicle for studying student processing behaviors in real instructional settings. Taking notes results in an external, reviewable record of the facts, concepts, relationships, and principles presented in the lecture. If notes mirror some of the cognitive processing activities which take place during listening, it can also be argued that notes provide a methodology to examine how students perceive and organize new information. As a result, the study of notetaking offers unique potential for exposing what is often hidden, that is, how students manipulate incoming information.

From an instructional practice perspective, notetaking is equally important to understand because it is a prevalent student behavior from entry to exit. In a typical lecture class, a large amount of student time is devoted to the recording, reorganizing, and reviewing of lecture notes. Whatever an instructor can do to improve the quality of the listening environment, including the use of clear and explicit lecture organization, highlighting techniques to emphasize critical information, as well as special accommodations for some students, should lead to more productive use of student time.

Recent research on notetaking has revealed interesting information about student processing activities during the presentation of information. Further work is needed to assess students' perspectives on the value of notes, how notetaking skills relate to different types of learning outcomes and the effects of direct training on the quality of notes and other study strategies.

Reference Notes


References


DiVesta, P.J., & Gray, S. Listening and note-taking: II. Immediate and delayed recall as function of variations in thematic continuity, notetaking and length of listening review intervals. Journal of Educational Psychology, 1973, 64, 276-278.


Norton, L. The effects of notetaking and subsequent use on long-term recall. Programmed Lear-

Journal of Instructional Development
Table 1
Notetaking and the Events of Instruction

<table>
<thead>
<tr>
<th>Instructional Event</th>
<th>A Lecturer Might Operationalize This Event in the Following Ways:</th>
<th>Which Encourages Notetakers to:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Gaining attention</td>
<td>Physical movement—c.g., Move to podium. Arrange lecture notes. Switch on overhead projector Verbal interaction—Good morning. Let's begin.</td>
<td>Prepare to listen&lt;br&gt;Locate place to take notes</td>
</tr>
</tbody>
</table>
| 2. Informing learner of objective         | Overview the lecture content  
Suggest how the information will be useful | Formulate a rationale why material should be noted  
Retrieve critical information from long-term memory to working memory  
Search for past associations  
Review earlier notes |
| 3. Stimulating recall of prerequisite      | Review terminology  
Summarize main points from previous lecture  
Ask questions to determine if students recall key terminology, concepts, principles  
Provide 5 minutes for students to review notes from an earlier session | Retrieve critical information from long-term memory to working memory  
Search for past associations  
Review earlier notes |
| learning                                   |                                                                                                                               |                                                                                                 |
| 4. Presenting the stimulus material       | Speak at a comfortable pace  
Provide salient organization of lecture points  
Pause to allow for questions, clarification | Alternate between own words and lecturer's words  
Discriminate between essential and nonessential information  
Uses mnemonics |
| 5. Providing "learning guidance"          | Use verbal cues such as "Note the following," "This is important to remember," "Record this in your notes"  
Use blackboard, overhead transparencies judiciously to highlight major points  
Raise questions to test comprehension of an idea  
Provide an outline | Practice performance by overtly or covertly responding to questions/problems  
Highlight notes material relevant to responses  
Correct inaccuracies  
Attend to essential information |
| 6. Eliciting performance                   | During lecture, provide sample questions/problems similar to those which will be presented on an examination                | Rehearse notes content in preparation for examination  
Integrate new information into existing notes |
| 7. Providing feedback about performance   | Request other students to respond to student answers  
Models responses to questions/responses |                                                                                                 |
| correctness                                |                                                                                                                               |                                                                                                 |
| 8. Assessing the performance              | Encourage distributed and massed review of lecture content  
Present divergent examples, nonexamples, and problem situations during lectures  
Explicitly link information from previous lecturer with current lecture |                                                                                                 |
| 9. Enhancing retention and transfer       |                                                                                                                               |                                                                                                 |
The Number of Performance Assessments Necessary to Determine Competence

Robert L. Lathrop
Professor, Vocational Education
Florida State University
Tallahassee, FL 32306

Abstract. Previous approaches to the issues of misclassifications and test length in criterion (competency, domain)-referenced assessment have usually been based on binomial or on Bayesian probability models which are not directly applicable to most instruction/selection decisions. The logic of the Sequential Probability Ratio Test, however, exactly parallels competency or mastery decision problems. It does not require a priori knowledge of individual examinee's in information about prior performance of other groups. It allows the examiner to establish reasonable levels of acceptance and rejection errors and specifies the minimum number of testing trials necessary to make individual decisions within these limits. The procedure, although mathematically unfamiliar to many practitioners, lends itself to a graphic solution with the use of only two tables.

Background
Every day each of us makes countless decisions about others based on our observations of their behavior. For those of us in educational enterprises, many of these decisions are judgments about another person's level of competence. In recent years educators have been interested in an ordinal measurement scale which describes individuals as having, or not having, competence (mastery) of some particular skill or knowledge.

The underlying problem faced by an educator making competency decisions is to attempt to describe each examinee's true level of competence, that is, to correctly describe the individual as competent or not competent according to a predetermined criterion. Because we are dealing with fallible assessment procedures and human inconsistency, some degree of decision error or misclassification is inevitable. The issue addressed in this presentation is basically a sampling question: How many observations of an individual's behavior are necessary to provide a prespecified risk of an erroneous decision about competence?

Although a number of writers have discussed this subject, their approaches have generally not proved useful to practitioners (Millman, 1972, 1973; Novick and Lewis, 1974; Hambleton, et al., 1976, 1978). Rather than describe these approaches to determining test length (number of assessments), we will instead consider an alternative approach developed by Wald (1947) as a means of making quality control decisions. Wald's approach, known as sequential analysis, seems to fit competence assessment problems perfectly.

Consider a situation where an instructor wishes to make mastery decisions about individual learners. We will make the assumption that at any given point in time a particular learner either has mastery or does not have mastery, according to some predetermined criterion. The instructor's task is to make decisions about individual learners in such a way that over a large number of such decisions, (s)he makes no more than some prespecified number of misclassifications. The reader should bear in mind that for each individual we make only one decision (at a time) and therefore, we are either right or wrong in that particular choice. What we would like is a decision rule that allows us to control the risk of making incorrect decisions.

Consider Figure 1, which illustrates the possible outcomes for mastery decisions. On the left margin we assume that each individual either has mastery or...
does not have mastery according to our definition. Obviously we do not know for any individual which category is correct for him/her. We do, however, have the results from our testing, and it is on this basis that we make the decision shown on the baseline of Figure 1.

If the examinee has Mastery and we so classify him/her, we have made a correct decision. If, on the other hand, our assessment leads to the conclusion that (s)he does not have mastery, we have made an error of “rejection.” In hypothesis testing this type of error is referred to as Type I or Alpha Error.

Now consider the converse. Suppose that the individual does not have Mastery. If we conclude, based on our assessment, that the individual does not have mastery, we have made a correct decision. It is possible, however, that our assessment erroneously leads us to conclude that the individual does have mastery. In this case, we would have made an “acceptance” error. This type of error is customarily referred to as a Type II or Beta Error.

Although we speak of Type I and Type II errors (or misclassifications), such errors exist only in an inferential sense. We do not know the “true” mastery status of an individual. Comparisons of our decisions to “true” status, therefore, are always inferences. To the extent that we act on the basis of our decisions as if they were true, the “true” status of an individual may be academic. In a baseball game a called pitch is a “ball” or a “strike” depending on what the plate umpire says it is. If this is the mode of using assessment data to make instructional decisions, “true” status may be conceptually interesting, but of little operational value. As the expression goes, “They is what I says they is!”

In a broader sense, however, erroneous decisions always have consequences whether we acknowledge them or not. In a recreational game a bad call may affect the final score. In an employment decision a “bad call” may result in employing a noncompetent worker (or rejecting a competent one). In an instructional decision a misclassification may require a learner to remain at some level of instruction longer than necessary (or, conversely, allow a non-competent learner to advance prematurely). In the final analysis, all misclassifications have consequences. Whether a particular misclassification is important is determined by the seriousness of the consequences. In some instances, the consequences of a classification error are borne primarily by the individual being assessed and may be inconsequential to the institution making the error. In other cases, both the individual and the institution (organization, agency, etc.) making the decision bear the costs of erroneous decisions. There are always costs for misclassifications whether we are conscious of them or not.

On the assumption, then, that we wish to exercise some conscious control over misclassifications, we would prefer a decision strategy that allows us to assign error risks in advance and make our decisions within these limits. The Sequential Probability Ratio Test developed by Wald (1947) meets this requirement. This procedure makes no a priori assumption about the level of performance of an individual examinee, but rather depends on the principle that an individual’s level will emerge from a relatively small number of trials. As decision makers, we want a strategy that requires us to draw the minimum number of samples of behavior to categorize an individual’s performance within whatever risk limits we have determined are reasonable. Stated another way, a good decision making strategy will allow us to quickly identify an individual who is clearly above or below some prescribed performance standard. The Sequential Probability Ratio Test (SPRT) establishes boundary values based on preselected levels of alpha and beta. The assessment of the individual proceeds (sequentially) until the pattern of assessments crosses either the acceptance or rejection boundary values.

Although it has sometimes been mistakenly believed that the Sequential Probability Ratio Test requires the assumption of independent observations, such is not the case. Wald points out that the basic inequalities on which the test is based are equally valid for dependent observations so long as the probability is 1.0 that the procedure will eventually terminate (pp. 43-44). The fact that the procedure is valid for a very general class of situations allows us to use the technique in practical testing situations where assumptions of independence of observations would not be reasonable.

**Procedure**

Suppose we wish to determine whether or not an individual learner has reached criterion performance on a particular learning task. We decide that to be competent (have mastery, reach criterion,...) (s)he must attain 80 percent of the maximum possible score on a performance checklist. We also determine, based on our subjective analysis of the consequences, that we wish to hold the risk of allowing a student without mastery to proceed to the next task, to a probability of .10 (error of acceptance, beta = .10). Similarly, we decide we want to set the risk of mistakenly holding back a student who has mastery to a probability of one-in-five (error of rejection, alpha = .20). Having made these determinations we can proceed.

Although the (SPRT) can be presented in either tabular or graphic forms, the writer recommends the graphic approach for the non-mathematically oriented user.

The graphic solution involves plotting two straight boundary lines on an X-Y graph. The horizontal axis is labeled m or (trials) and is simply the number of sequential observations (assessments, tests, trials...) for a particular examinee. This number will vary from individual to individual. The vertical axis is labeled cm, and is the cumulative number of nonpassing trials for an individual examinee (See Figure 2).

The two sloping lines, labeled L_0 and L_0', represent the two boundary lines for our decisions. If we plot m and cm for
any individual examinee, the plotted points will always begin between the base line, and $L_y$. As the testing proceeds the plotted points for each examinee will eventually cross either $L_x$ or $L_y$. As soon as the plotted points cross either boundary line, the testing is complete and an appropriate decision made.

Because $L_x$ and $L_y$ are straight parallel lines we need to calculate only three values: the point where $L_x$ crosses the vertical axis, the point where $L_y$ crosses the vertical axis, and the slope of $L_x$ and $L_y$. Since $L_x$ is the rejection (nonmastery) boundary line, we will label this intercept $r_x$. Similarly, the point where $L_y$ crosses the vertical axis is the acceptance (mastery) intercept which we will label $a_y$. Since the lines are parallel, they will both have the same slope, labeled $s$. To establish line $L_y$ on our graph, we need to plot two points. We will plot the intercept (where $m = 0$) and the value for $m = 10$.

The mathematical expressions for $r_x$, $a_y$, and $s$ are shown by Wald to be as follows:

$$r_x = \frac{\log_{10} \left( \frac{1 - \beta}{\alpha} \right)}{P_0 (1 - P_1)}$$

$$a_y = \frac{\log_{10} \left( \frac{\beta}{1 - \alpha} \right)}{P_0 (1 - P_1)}$$

$$s = \frac{\log_{10} \left( \frac{P_0 (1 - P_1)}{1 - P_1} \right)}{1}$$

Where:

$\alpha$ = rejection error level
$\beta$ = acceptance error level
$1 - P_0$ = upper limit of criterion tolerance (indifference) region
$1 - P_1$ = lower limit of criterion

Although these expressions may appear somewhat intimidating, fortunately they lend themselves to rather simple solutions which can be tabulated for commonly used values.

Before proceeding with our example, the preceding mathematical expressions contain two terms, $P_0$ and $P_1$, which deserve further comment.

The context of Wald's work was the need to make sampling decisions about the ratio of defective to nondefective parts in industrial shipments. Wald suggested that in most practical situations the decision maker was less interested in knowing the "true" ratio of defects in a shipment than (s)he was setting some upper and/or lower boundary for the "true" percentage. That is, that $P$, the "true" percentage of defects is no greater than $P_0$ nor less than $P_1$. As long as $P_0 < P < P_1$, Wald stated that the decision maker was indifferent to the "true" value of $P$. Within the "zone of indifference," to use Wald's term, there is no practical difference between $P$ and $P_0$ or $P_1$.

Obviously the choices of values for $P_0$ and $P_1$ are judgmental matters which go beyond purely statistical considerations. In the context of performance testing, if we set mastery at 80 percent, $P$ becomes 20 percent. That is, we will allow 20 percent "defective" performances and still accept the examinee as competent (having mastery). The examinee then has a choice of how closely (s)he wishes to hold to exactly 80 percent. Is the distinction between 79 percent and 80 percent meaningful (or, in terms of $P$ and $P_1$, is the difference between 21 percent and 20 percent significant)? In a more general sense, how far can $P_0$ and $P_1$ range away from $P$ before the difference has practical significance? In most educational testing the necessity to make extremely precise estimates of the "true" value of $P$ for an individual is less important than the ability to quickly sort the examinees into three broad categories:

1. Those individuals who clearly are well above the minimum requirements for mastery;
2. Those individuals who are well below the minimum requirement for mastery;
3. Those individuals who require additional testing before a mastery determination can be made.

The more precision we require (that is, the closer we set $P_0$ and $P_1$ to $P$), the larger must be the number of testing trials. In the examples and tables developed for this presentation, the writer has set $P_0 = P - .10$ and $P_1 = P + .10$. Translated into the context of this example if:

$$P = .20$$
$$1 - P = .80$$

Nominal criterion level for mastery

$$P_0 = .10$$
$$1 - P_0 = .90$$

Upper limit of Region of Indecision (Zone of Indifference)

$$P_1 = .30$$
$$1 - P_1 = .70$$

Lower limit of Region of Indecision (Zone of Indifference)

By setting $10 \times P < .30$, we have, in effect, said we want to quickly identify examinees who can perform acceptably on at least 70% of the trials (with an accompanying misclassification risk = alpha). The remaining group whose performance is above 70% but less than 90% we say is "too close to call" without further testing. The practical question the examiner must then ask and answer is, "Does the decision justify additional testing or am I willing to make a decision based on the information available, taking into account the seriousness of the consequences of misclassification?" In most performance testing problems familiar to the writer, the values proposed for $P_0$ and $P_1$ are a reasonable compromise. The Sequential Probability Ratio Test procedure allows the user to set whatever values for $P_0$ and $P_1$ (s)he chooses. The tables presented here, however, apply only to $P_0 = P - .10$ and $P_1 = P + .10$.

Table 1 presents intercept values for $r_x$ and $a_y$ for several combinations of alpha and beta. The values in Table 1 are based on a criterion level of 1 - $P = .80$. Because decisions for individuals having mastery levels slightly above or slightly below .80 will be very difficult to determine, we will establish a set of tolerance limits around 1 - $P$ within which we say decisions are not practically possible. In Table 1 the tolerance limits have been chosen as (1 - $P$) ± .1, making 1 - $P = .79$ and 1 - $P = .89$. From Table 1 we locate intercept values $r_x$ and $a_y$ for $P = .20$; beta = .10.

We also find from Table 1 that

$$s = 1.86$$

The values for $r_x$ and $a_y$ are plotted on the vertical axis of Figure 2 where $m = 0$. Because of the ease of calculation we will calculate the values of $L_x$ and $L_y$ when $m = 10$.

$$L_x = r_x + m(s) = 1.11 + 10(1.86) = 2.67$$

$$L_y = a_y + m(s) = -.54 + 10(1.86) = 13.32$$
Plotting these values we have the lines $L_1$ and $L_0$, shown in Figure 2.

We are now ready to begin our sequential testing. We administer our assessment procedure to Student "A." One possible outcome of this assessment is "pass." if the student meets or exceeds the 80 percent criterion. In this case his/her dm score is 0. If the student does not reach criterion on this trial his/her dm score = 1. No decision about mastery is possible on the first trial since both dm = 0, and dm = 1, fall between $L_1$ and $L_0$.

We proceed to trial number two ($m = 2$). The possible cumulative outcomes are:

- $dm = 0$, if the student reaches criterion on both trials
- $dm = 1$, if the student reaches criterion on one trial
- $dm = 2$, if the student reaches criterion on neither trial

Since $dm = 2$ for trial two falls above $L_1$, we reach a decision of non-mastery for any student with a dm score of 2 on the second trial. Students with scores of $dm = 0$ or $dm = 1$ on the second trial would continue to be tested.

The sequential testing continues either until the plotting of dm scores for an examinee falls outside $L_1$ or $L_0$ or the examiner decides to terminate the testing because of practical limitations. It can be shown, theoretically, that every dm plotting will eventually cross either $L_1$ or $L_0$. In practice, however, this number of trials may not be justified. The failure of a dm plot to quickly move out of the region of indecision and cross $L_1$ or $L_0$ indicates either that the examinee's true level of performance is very close to the criterion level, or that the assessment procedure is somewhat unstable (unreliable).

Further examination of Figure 2 reveals that at this criterion level ($1 - P_0 = .8$), decisions involving mastery (crossing $L_0$) require more trials, than decisions of nonmastery (crossing $L_1$). Note that, not until an examinee has performed at or above criterion levels for nine successive trials would his/her dm plot cross line $L_0$. On the other hand, for an individual with as few as two unsuccessful trials, we could make a nonmastery decision.

From Table 1 we can see that if we require alpha or beta to be small (low risk of a misclassification), the intercepts and therefore lines $L_1$ and $L_0$ are further apart. The effect of decreasing the acceptable error limits, therefore, is to increase the number of trials (assessments) needed to make a decision. Conversely, we would set alpha as low as practical. We might in this case allow beta (labeling a nonmaster as a "master") to go to .5, the level of chance. Plotting the intercepts for alpha = .2 and beta = .5 on Figure 3, we see that with as few as three trials, a mastery decision could be made (solid lines $L_1$ and $L_0$).

If, in this example, we had been primarily concerned with not labeling a person with mastery as a "nonmaster," we

Table 1
Slope and Intercept Values $r_1$ and $a_0$ for Various Levels of $\alpha$ and $\beta$
Where $1 - P_0 = .9$; $1 - P_1 = .7$

<table>
<thead>
<tr>
<th>$a_0$</th>
<th>.1</th>
<th>.2</th>
<th>.3</th>
<th>.4</th>
<th>.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.63</td>
<td>1.11</td>
<td>.514</td>
<td>.601</td>
<td>.435</td>
</tr>
<tr>
<td>.1</td>
<td>-1.11</td>
<td>.514</td>
<td>.727</td>
<td>.513</td>
<td>.348</td>
</tr>
<tr>
<td>.2</td>
<td>1.54</td>
<td>1.03</td>
<td>.727</td>
<td>.513</td>
<td>.348</td>
</tr>
<tr>
<td>.3</td>
<td>.144</td>
<td>.928</td>
<td>.628</td>
<td>.415</td>
<td>.249</td>
</tr>
<tr>
<td>.4</td>
<td>.601</td>
<td>.513</td>
<td>.415</td>
<td>.300</td>
<td>.135</td>
</tr>
<tr>
<td>.5</td>
<td>.19</td>
<td>.679</td>
<td>.378</td>
<td>.249</td>
<td>.135</td>
</tr>
</tbody>
</table>

Slope = .186

Figure 2. Graphic solution to the sequential probability ratio test for $\alpha = .2$; $\beta = .1$; $1 - P_0 = .9$; $1 - P_1 = .7$
relatively indifferent to overlooking some individuals with mastery (say, in a selection situation where we have more qualified applicants than positions), we could set \( \beta = .1 \) and allow \( \alpha \) to operate at a chance level (\( \alpha = .5 \)). In this case, plotting \( L_a \) and \( L_b \) (dotted lines) for \( \beta = .1 \) and \( \alpha = .5 \), we find that we can reach some nonmastery decisions on the first trial, but would not make any mastery decisions with less than seven successive passing trials (\( dm = 0; m = 7 \)).

In each of the above examples we have used a criterion standard of .80 with a tolerance region of plus or minus .10. Changing the criterion standard for mastery from .80 to another value will affect both the slope and the intercept of \( L_a \) and \( L_b \). Although separate tables parallel to Table 1 could be constructed for each possible criterion-level, the relationship between criterion level and slope and criterion level and intercept can easily be shown in a second table (See Table 2).

Consider an example where the criterion for mastery was .70 instead of .80. To compute intercepts \( r_1 \) and \( q_0 \), we extract the appropriate values from Table 1 for our selected levels of alpha and beta and multiply these values from Table 1 by the intercept weight shown in Table 2. For example, if we set \( \alpha = .20 \) and \( \beta = .30 \), the values from Table 1 are \( q_0 = -0.727 \) and \( r_1 = .928 \). If we multiply each of these values by 1.38 we have \( q_0 = -1.00 \) and \( r_1 = 1.28 \). The slope is read directly in Table 2 as, \( s = .293 \).

We can now plot \( L_a \) and \( L_b \) as shown in Figure 4. For comparative purposes \( L_a \) and \( L_b \), corresponding to a criterion level of .80, will be shown as dotted lines in Figure 4.

**Summary**

Previous approaches to the issues of misclassifications and test length in criterion (competency, domain)-referenced assessment have usually been based on binomial or on Bayesian probability models which are not directly applicable to most instruction/selection decisions.

The logic of the Sequential Probability Ratio Test parallels, exactly, the logic of most competency or mastery decision problems. It does not require a priori knowledge of individual examinees nor information about prior performance of other groups. It allows the examiner to establish reasonable levels of acceptance and rejection errors and specifies the minimum number of testing trials to make an individual decision within these limits. The procedure, although mathematically unfamiliar to many practitioners lends itself to a graphic solution with the use of only two tables. Once plotted, the selection consequences of various choices for alpha, beta, criterion level (\( 1 - P \)), and tolerance limits become obvious.

The practitioner may find several of the following conclusions helpful in his/her use of the Sequential Probability Ratio Test to make instructional/placement/selection decisions regarding mastery.

1. The choice of levels for alpha and beta should be based on realistic estimates of the consequences of a particular misclassification. In some selection problems the "cost" of an error may fall primarily on the institution doing the selection as in the case where an applicant who is not qualified is selected. In others the "cost" may be borne entirely by the qualified examinee who is not selected. Both individual and institutional costs should be considered in setting alpha and beta levels.

2. In all cases, alpha and beta should be allowed to be as large as practical, if
decisions are to be made in the minimum number of trials. If alpha and beta must be kept at very low levels, the number of trials to make such decisions must be expected to increase accordingly.

3. In many testing problems we are interested in controlling only one type of error. If, for example, we are screening individuals on a pretest, we may be relatively indifferent to making a Type I error (not exempting a student who already has mastery) on the argument that even the student who has minimal mastery could benefit from some "overlearning." In selection decisions examiners are typically interested in minimizing Type II errors (erroneously selecting a nonqualified candidate, admitting a nonqualified student, etc.). Where classification decisions can be revised based on subsequent observations, the consequences of either a Type I or a Type II error can be minimized and therefore initially allowed to remain somewhat large.

4. Having established levels for alpha and beta, the number of testing trials becomes a function of the tolerance limits for the region of indecision, \(1 - P_o\) and \(1 - P_i\). If the nature of our decision requires us to make discriminations close to the specified criterion level, that is, the values \(1 - P_o\) and \(1 - P_i\) are set numerically close to the criterion level \(1 - P\), then we must expect to collect a large sample of test behavior before reaching a decision. Conversely, if we are willing to classify individuals somewhat broadly and are willing to suspend decisions for persons with test results close to the criterion, we can reduce the number of testing trials substantially. In all cases we should allow the range of \((1 - P_o) \times (1 - P_i)\) to be as wide as possible.

5. In virtually all cases, more than one trial or performance assessment is required to reach a decision about mastery. Under certain circumstances a decision that an individual does not have mastery can be reached in only a few trials, perhaps two or three. In contrast, rarely would we be able to decide that an individual had mastery with less than four or five successful trials.

6. There is no way in advance to determine how many trials will be required to make a decision about an individual examinee. The process, however, can be discontinued at the end of some predetermined number of trials, with those examinees for whom a clear pattern of performances has not emerged being placed with whichever decision has the less serious consequences.

7. The number of trials necessary to make a decision is influenced by the choice of the criterion-level. In general, the effect of lowering the criterion-level is to increase the number of trials needed to make a decision. A further consequence of lowering the criterion-level is to increase the slope of the decision limit lines. As the slope of the decision limit lines increases, the number of trials required for nonmastery decisions increases and the number for mastery decisions decreases. The reasonableness of this consequence is intuitively obvious; as we lower the level of performance required for mastery we reject fewer persons for lack of competence and accept more persons as having mastery.

In all cases the Sequential Probability Ratio Test provides the smallest number of sample observations necessary to make binomial decisions within predefined error limits. In many cases reliable decisions can be made with half the number of observations suggested by other sampling procedures (Wald, 1947). Because a smaller number of undecided cases remains after each successive testing, the Sequential Probability Ratio Test allows us to quickly and efficiently separate the individuals with consistently high or consistently low performance from those individuals whose performance pattern is not clear-cut. For a given expenditure of testing effort, Cronbach and Gleser (1957) have shown sequential sampling as being the most efficient of several decision strategies compared. The procedure outlined in this presentation will allow this powerful decision-making technique to be used by educational practitioners for a variety of selection/placement decisions.

References


Just about everyone seems to be considering a job or career change these days. Whether seriously looking or just exploring options, a reader will find any number of books on the market promising to provide guidance and moral support during the search. In reality, no one book thoroughly covers all aspects of the subject. However, each of the titles reviewed here may provide job and career changers with some of the help they need.

What Color is Your Parachute? is something of a classic in the career counseling field. Upbeat and humorous as well as informative, the book guides the reader through the process of clarifying career goals, identifying target organizations, and making contact with the people in those organizations who do the hiring. Bolles includes an extensive bibliography of specialized information sources, an essay about coping with burnout, and a discussion of the rewards and pitfalls of seeking help from professional counselors.

The instructional developer thrashing about in search of a more suitable line of work will find Bolles' exercises on identifying career aspirations alone to be worth the price of the book. Similarly, the person trying to break into the field will profit from Bolles' advice on establishing a network of contacts through informational interviews. But for the instructional developer with a number of years and acquaintances in the field, Parachute may be of limited value beyond providing general background about the job market.

For the individual who has progressed beyond the career exploration stage, Lathrop's Who's Hiring Who might be a better choice. Richard Lathrop describes the nation's job market as "a chaotic mess." and job hunting as an acquired skill. He goes on to teach that skill through a series of well-designed exercises and performance aids.

Lathrop focuses on the practical side of job hunting. He begins by suggesting a self-examination session—this time for the purpose of identifying the reader's strongest selling points. Then Lathrop demonstrates how to take those selling points and hammer them into an effective resume, which he calls a qualifications brief, because it summarizes accomplishments rather than previously held jobs. Lathrop even supplies a formula for writing cover letters which almost demand a potential employer's attention. Next, he explores techniques for preparing for interviews, and provides lists of typical interview questions, with "right" answers. For those readers not ready to change jobs, there is a chapter entitled "Double Your Pay with Your Present Employer."

The Professional Job Changing System, by Robert Jameson Gerber, is a useful source for any professional person, whether involved in job changing or job hunting for the first time. Gerber's suggestions are sometimes aimed just at entry level workers, sometimes just at middle or upper management. Such suggestions are clearly differentiated by audience. Instructional developers will benefit by reading all the suggestions, and making changes in sample answers to questions and the like to suit the professional seeking employment at any level.

Gerber is some of the same ideas as Lathrop, but he includes many more tips on developing interview opportunities. Techniques covered include making and using personal contacts, answering and placing advertisements, mounting a direct mail campaign, using employment agencies, executive search firms, and placement counselors, and using dramatic approaches like making cold telephone calls and sending telegrams and registered letters.

Gerber talks about writing resumes and letters, and includes lots of samples. Also included is a proposed way of developing one's personal data base. Most useful is a chapter on overcoming liabilities, like not having enough experience or the right experience, or the right degree, or being too old, needing to change careers, or being unemployed. A chapter for women has helpful ideas for anyone who wants to handle personal questions gracefully, or needs to become more aggressive.

For those who want help just with handling interviews, Medley's Sweaty Palms is recommended. Seasoned instructional developers and those coming into the field with previous work experience will need to overlook suggestions aimed at new college graduates, but everyone will find Sweaty Palms a real help in getting through interviews.

Medley's theme is that the interviewee must concentrate on the interviewer, and help the interviewer make the process go smoothly. It is the interviewee's responsibility to control the content of the interview—so the qualifications and accomplishments of the interviewee stand out in the interviewer's mind. The interviewer is trying to formulate an im-
prescription of the candidate, and it is the candidate's responsibility to make the impression a favorable one.

Medley includes hints on preparing for interviews, showing enthusiasm answering questions, acting on assumptions about the interviewer, honesty, confidence, nervousness, dress, the use of silence, and salary negotiations. Appendices include lists of commonly asked interview questions, with more "right" answers. The style is informal and slightly wordy, and some of the suggestions are repeated in more than one chapter. The chapter entitled "Sex" might better have been labelled "for women," since it does not treat fully the problem encountered by men in answering personal questions.

Instructional developers, who must have interviews with employers in a wide range of fields and types of organizations, will find the chapter on preparing for interviews especially useful.

All four titles maintain the positive, upbeat approach everyone needs when looking for a job. Instructional developers, whether earnestly looking for that first job, or seeking a new position, must keep up their self-confidence and perseverance. All four titles offer real assistance in what should be regarded as a challenging, rewarding experience—looking for a job.—Revised by Ann L. Wiley, Ann L. Wiley Consultants, Auburn, NY. and Helene Geiger, Rochester, NY.
motivation, syntax, symbolism, etc.) is often truncated because viewers apply shallow (automatic) processing schemata which they deem appropriate to "entertainment" or casual dissemination of information. Salomon argues that most viewers tend to avoid using deeper (non-automatic) schemata unless they perceive messages as being clearly "educational."

A sample application of Salomon's use of AIME to explain media effects: learners who have acquired a perception of television as a "shallow" medium from which information is easily acquired will manifest a low AIME (and learn little) even if the material is information-rich. Salomon cites studies which suggest that learners confuse ease of learning with amount of learning. It appears that learners who feel that learning from television is easier than learning from books, actually learn more from books in controlled studies.

Salomon's review of this and similar findings suggests that instructional developers must consider the effect that tens of thousands of hours of passive television viewing has on learner attitudes towards instructional media. Has television impaired the ability of learners to acquire knowledge and skills from instructional videotapes and films and tape/slide presentations by reducing the mental effort they are willing to devote to "TV-like" media? Will learners expect new communications media such as interactive video and educational computers to be entertaining and shallow and thus shape (through market demand) these media to confirm their expectations. (A common complaint of media producers is that contemporary audiences evidence a constantly accelerating demand for presentations that are "entertaining"—that is, presentations that incorporate special effects that are ever more sophisticated and intense.) Communication and Education represents a major step in addressing these issues.

As one considers the implications of Salomon's ideas for the future of educational media, it is hard to overlook the current popularity of videogames, which take the biggest share of consumer dollars spent on computers. Obviously, videogames place great emphasis on immediate gratification. But Salomon's discussion suggests that the most salient characteristic of this medium may be its requirement for automaticity in processing information. Performance in the world of video arcades seems to place great importance on rapid reduction in the amount of invested mental effort. Are videogames now implanting schemata which will further reduce the mental effort that masses of computer users will be willing (or able) to devote to the interactive computer-based technologies of the future? It is a question worth pondering. Remember Dr. Seuss? The 78-year old author of over 50 children's books just signed an exclusive contract with a major manufacturer of video games.—Reviewed by Brockenbrough S. Allen, Assistant Professor of Educational Technology, San Diego State University.


Judging by the topics found most often in the literature on instructional design, we seem to have been most successful in documenting the "systems approach" as a whole. We are also making progress toward increasingly precise methods of front-end analysis and evaluation of learning outcomes. However, the actual selection of instructional strategies at a detailed level and development of finished materials seems to have received the least formal attention.

Designs for Instructional Designers addresses the need for an informed way of translating a front-end analysis into completed instruction. The book emphasizes heuristics, not procedures. Consequently, there is still ample room for intuition and creativity, but the book should help students and experienced developers alike improve their judgment about the appropriateness of given instruction in a given situation.

Although the book explicitly deals with programmed instruction, the principles discussed apply universally. In Chapter One, three basic principles of instructional design are presented: active responding (which Markle suggests ought to be referred to as "meaningful responding"), "errorless" learning (including a nicely balanced discussion of "lean" programming), and immediate feedback. Then follows a detailed examination of a segment of programmed text in which issues are raised such as the relationship of objectives to instruction, assumptions about prerequisites, and details of design including use of positive and negative examples, questions, and concept-level sequence. The discussion refers to principles of concept analysis, rather than being limited to a strictly objectives-based approach.

Chapter Two picks up the theme of "lean" programming by discussing the discrimination of need-to-know from nice-to-know and the role of task and audience analysis in determining what to teach. The chapter concludes with some useful advice on the instructionally effective use of typography and layout.

The third chapter deals with sequencing of instruction, both at the levels of individual concepts and entire skills. This includes a discussion of priming and prompting (for concept teaching), and use of learning hierarchies, concept hierarchies and algorithms (for skill sequencing).

The final chapter summarizes two well-known attempts at more refined procedures for sequencing of concept instruction: Gilbert's Mathematics system and Evans, Homme, & Glasser's RULIEG system. Both presentations are detailed enough for the learner to actually use them, but both are presented as only part of the truth, rather than as any final answer.

The book is in a large-frame quasi-programmed style which might more accurately be described as a text which uses inserted questions. The tight learner management and testing usually associated with programmed instruction are not included, nor is any explicit branching. However, headings, layout, and typography are used in instructionally meaningful ways, so a fair amount of learner control of instructional sequence is possible. For example, digressions on research evidence for various points are isolated from the main text, to be read at the learner's discretion. In short, unlike many authors on instructional design, Markle did follow her own advice.

Markle's overall tone in the book is refreshingly pragmatic. Numerous positive and negative examples are used, even when they are quite long. The writing style is informal, and ideological exhortations are kept to a minimum. In keeping with the "heuristic" approach, the general impression is that while the information presented will help the learner refine his or her judgment, general procedural/prescriptive synthesis is not an objective of the book.

Also refreshing is the theoretical eclecticism of the book. I have long believed that practitioners of instructional design must be theoretically grounded in what they do (where possi-

If you are looking for some yummy pleasure reading for an upcoming long weekend, Romiszowski’s Designing Instructional Systems is absolutely the wrong choice for you. If, on the other hand, you are willing to work hard, and come back repeatedly with questions and comparisons, then this comprehensive professional text is an excellent choice. While it is ugly, structurally complex, and occasionally redundant, it is also detailed, significant, integrated, international, and unique in its ability to explain the principles, concepts, and procedures of instructional systems development and their theoretical underpinnings. The text costs $32.50 and is worth it. It is an important addition to a professional library.

Designing Instructional Systems is about front-end analysis. The two major sections of the book are defining and analyzing the problem. The research, debate, strategies, and options for approaching large scale educational or training problems are definitely here—in detail, with appropriate allusions to significant contributors, and with usefully drawn comparisons. Whether he intends to or not, Romiszowski is writing the book for an in-service audience, readers who will pick it up to see what he and others have to say about objective hierarchies, or task analyses, or inductive vs. deductive strategies. That kind of a specific need will yield historical, theoretical, and procedural details and levels of information. The very same structural complexity and unbroken comprehensive prose which renders the book of so much interest to a professional would make it unpopular with the student of instructional design. This handy “tool box” (Romiszowski’s words) of multiple procedures for experienced developers would probably annoy and baffle any but the most highly motivated neophyte.

This text has numerous strengths:
1. It is an important step towards uniformity in our professional language. Romiszowski presents useful and convincing distinctions between, for example, education and training, macro and micro problems, and performance and information as instructional intentions.
2. The instructional theories of Gagne, Gilbert, Skinner, Bruner, Mager, Pask, etc., are compared and examined for their implications for instructional design.
3. The British perspective is refreshing and important. The author is much more conversant with our researchers and theorists than most Americans are with the work of Seymour, Wyant, Rowntree, or Rackham.
4. Romiszowski is pleasingly full of opinions. For example, he charges Gagne with “woolly conditions” and laments ISD failures to solve social and personal performance problems.

The weaknesses of the book come, perhaps, from the absence of irreverence. Romiszowski loves the tools of our trade and he provides the reader with innumerable charts, graphics, boxes, and flowcharts. Many of them are useful (e.g., for subject matter and task analysis examples); others, however, attempt to sanctify personal opinions by putting them in neat boxes (e.g., grouping considerations).

Romiszowski chunks his book and names the chunks in tiny side subheads. The chunks are too small, their placement is distracting and they add to the impression that the organization of the book is confusing rather than enlightening. It is sometimes confusing; it is also almost always worth the struggle to figure it out.

While Designing Instructional Systems fails to model ISD principles, it does effectively describe, compare, and contrast them. Analysis, design, and development receive effective, comprehensive treatment. A promised second volume is likely to address implementation and evaluation with the same care. You will not enjoy this book; you will, however, be very glad you are able to refer to it as you solve instructional problems.—Reviewed by Allison Rossett Professor of Educational Technology, San Diego State University.

Developed by Litton Industries for use in analyzing training problems and designing appropriate systems of individualized instruction to address those problems, the Instructional Systems Analysis and Selection (ISAS) procedures include questions, comparison matrices, guidelines, and suggestions designed to assist in the analysis of specific problems. The objectives of the analysis are to choose an instructional systems model that best fits the training need and to develop the highest quality and most cost-effective instructional materials possible within that model. A description of the overall system for delivering instruction, such a model describes the flow of trainees through the system, indicates the important decision points, and makes it possible to predict how the system will behave under different conditions. Several possible models are proposed and discussed in this paper.—Microfiche 97 cents, paper copy $2.00 plus shipping as document ED 200 770.

Instructional Design Process for a Course on Time Management for Head Nurses and Supervisors at a Veterans Administration Medical Center, Adrian D. Geering. Adelaide, Australia: Adelaide College of the Arts and Education, 1980. 97 pp.

Developed for a course conducted at the Veterans’ Administration Medical Center in Lincoln, Nebraska, the instructional design process described in this paper is based on A New Instructional Development Process for Instructors of Adults by Mary Jane Even. The paper covers course organization, course planning, instructional planning, analysis of the instructional plan, and evaluation of the course. Consideration of seven essential components is built into its design: the instructor, the adult learner, the teaching-learning process, the subject matter content, the group characteristics, the climate (environmental setting), and internal and external influences on the group. The paper contains a bibliography, needs assessment survey, content need assessment, course workbook, conference evaluation form, evaluation of objectives, personnel job descriptions, and a proposal for management development to be submitted to the Cooperative Health Manpower Education Program.—Microfiche 97 cents, paper copy $6.95 plus shipping as document ED 195 855.


This report describes the instructional design for a Multiple Position Letter Sorting Machine (MPLSM) operator’s course administered via the Automated Instructional Development System (AIDS). AIDS is an innovative system consisting of a unique special purpose computer, innovatively designed software, and individualized adaptive instruction. The design and specifications...
for the course are described in six sections and two appendices. Included are descriptions of the pedagogical principles which form the basis for the course design, course structure, course organization, remediation strategies, student feedback modes, refresher training, and the roles of the instructor. Appendices detail the module objectives and entry characteristics.—Microfiche 97 cents, copy $3.65 plus shipping as document ED 186 016.


Instructional Systems Development (ISD) is a systems approach to curriculum development and instructional delivery which is oriented toward occupational needs, with emphasis on what it is that students must learn to perform specific tasks, what facilities provide the best setting for the necessary learning, and what instructional methods and media best facilitate learning. There are four important differences between ISD and more traditional forms of instruction: (1) the requirement for the thoughtful selection of what is to be taught based on data from the field, e.g., job analyses, occupational surveys, and feedback from graduates; (2) attention given to how training is to be conducted, especially the consideration of innovative alternatives; (3) the use of test data based on absolute standards of performance to grade students and judge quality instruction; and (4) the application of contemporary technology to optimize instructional effectiveness, efficiency, and cost. Following the narrative discussion, the report describes the five phases of ISD—analysis, design, development, implementation, and control/evaluation—with charts showing the sequence of events. Event outcomes are also listed, and a six-page glossary is appended.—Microfiche 97 cents, paper copy $2.00 plus shipping as document ED 194 805.


Training provided in the Basic Medical Specialist course, which has one of the largest enrollments of the U.S. Army's Academy of Health Sciences, encompasses both emergency first aid (for field medics) and basic nursing skills. A task force working to improve Army training developed this course in accordance with the Interservice Procedures for Instructional Systems Development (ISD) model. The initial task analysis resulted in the separation of essential skills from those which were desirable but not necessary. For each task included in the course, instructors received administrative instructions, a lesson plan, a laboratory checklist, and standards for judging task performance. Each lesson was presented on videotape and students practiced the skills, generally using the buddy system, until mastery was achieved. Evaluation included a pretest and a posttest for each module. Preliminary evaluation of this course indicated that participant attitudes were favorable, student attrition was very low, and costs were reduced as compared to previous standards is appended.—Microfiche 97 cents, paper copy $3.30 plus shipping as document ED 190 628.


Intended for use by persons in developing countries responsible for initiating or expanding the use of audiovisual facilities and techniques in industry, this manual is designed for those who have a limited background in audiovisuals but need detailed information about how certain techniques may be employed in an economical, efficient way. Part one, which focuses on the use of audiovisuals, includes a brief introduction and discussions of factors involved in the selection process and in the actual use of such materials, including cost effectiveness and facilities. Under the heading Techniques and Tools, the second part presents some basic considerations in preparing art work for projected materials; brief descriptions of such non-projected aids as storyboards, flipboards, and flannelboards, together with their advantages and disadvantages, construction, and advice on use; general instructions for voice production: equipment for projected still pictures and guidelines for producing effective presentations; information on motion picture projectors and production techniques; and an overview of the types of videotape recorders available with suggestions for using VTR equipment and closed-circuit television. Procedural diagrams, pictures, and figures supplement the written text, and an internal directory of institutions and organizations concerned with film making or educational media is attached.—United Nations Industrial Development Organization, P.O. Box 300, A-1400, Vienna, Austria (free), or microfiche 97 cents, paper copy $6.95 plus shipping as document ED 192 711.


This paper describes competency-based systems and systems components in vocational and occupational education and other areas of elementary, secondary, and postsecondary education. Seven systems of competency-based instruction are discussed: Program for Learning in Accordance with Needs, Individually Prescribed Instruction, Individually Guided Education, PLATO (Programmed Logic for Automated Teaching Operations), Computer Curriculum Corporation's CCC 17 Systems, Bell Lab's Individualized and Group Instruction Systems, and Oakland Community College Cognitive Style Mapping Program. Three newly emerging systems are also described: Individualized Mathematical Systems, Kettering Secondary School Improvement Project and Challenge Education. Seven career guidance and information systems described are Planning Career Goals; AIR Career Guidance Program; Information System for Vocational Decisions; Discover Foundation, Inc.; CHOICES; Career Planning and Support System; and System of Interactive Guidance and Information. Curriculum materials in occupational and vocational education that are examined include those developed by the Mid-American Vocational Education Curriculum Consortium, Vocational Technical Education Consortium of States, and eleven other educational institutions or agencies. Assessment systems described are the California Test Bureau's (CTB) ORBIT (Objective Referenced Bank of Items and Tests), Houghton Mifflin's SCORE, Educational Testing Service's Program for Assessing Youth Employment Skills, and NOCTI (National Occupational Competency Testing Institute). CTB's
data management system, TRACER, is also described. — Microfiche 97 cents, paper copy $3.65 plus shipping as document ED 204 624.


This module on elements and procedures involved in planning instruction is the fifth of eleven modules in the set, Introduction to Teaching Adults. Designed to meet the learning needs of part-time continuing education instructors, these modules can be used as resource materials for local workshops or study-discussion groups, as self-instruction (each module takes approximately two hours), and as a correspondence course. (Source information on correspondence study is cited in the first module, ED 208 201.) All modules except the first are to be used in conjunction with an audiotape. The objective of the module is to enable the teacher to identify content for a course, arrange it in a logical order, and prepare lesson plans for use in teaching adults. The booklet first discusses some factors that affect decisions about what to teach and how to teach it. It suggests a procedure to follow in designing instruction that includes four basic phases: specifying objectives, identifying and arranging learning tasks, selecting techniques and devices, and developing evaluation instruments. A case study is presented to illustrate the application of the process. — University of British Columbia, Centre for Continuing Education, 5997 Iona Drive, Vancouver, BC V6T 2A4 ($40.00 per set, including cassettes for Canadian residents; $110.00, plus postage and handling, for international orders), or microfiche 97 cents, plus postage as document ED 208 205.


A study was conducted to test the hypothesis that Brunerian learning theory can provide the instructional designer with a framework for developing effective learning materials. To determine three levels of spatial ability, two standardized tests—the Spatial Visualization Test (SVT) of the Dayle Vocational Tests and part VI of the Guilford-Zimmerman Aptitude Survey—were administered to 34 students in the introductory statistics course of the educational technology graduate program at Concordia University. The high, medium, and low spatial ability subjects were randomly assigned to one of two treatment groups: an instructional simulation developed according to Brunerian learning theory and a traditional textbook approach. The two groups were administered a pretest, an immediate posttest, a one week delayed posttest, and a five week delayed posttest. The first three tests each consisted of five problems with three questions per problem. A 2 Treatment x 3 Spatial Visualization x 3 Test Position ANOVA was performed on the raw scores for three achievement tests, and the means and standard deviations were calculated. No sex differences were found in spatial ability, but the Brunerian simulation was found to be significantly beneficial for low spatial ability students. Findings demonstrated the usefulness of isolating the significant aptitude required for a specific learning task, as well as the relevance of Brunerian theory for instructional design. Five figures, six tables, and 58 references are provided.—Microfiche 97 cents, paper copy $3.30 plus shipping as document ED 217 856.


In order to investigate the effects of sequence and synthesis in the teaching of taxonomically-related concepts, a study was conducted in which 27 students from Syracuse University were asked to examine printed instruction dealing with kinds of sailboats, and then to respond to a test based on those instructions. The synthesizing structure employed in the instructions was a "kinds-conceptual" taxonomy which shows the relationship between concepts. Six versions of the instructions were employed: three with a general-to-detailed presentation sequence and three with a detailed-to-general sequence. Each set of three included a version without a synthesizer, a version with a synthesizer at the beginning, and a version with a synthesizer at the end. Statistical analysis was performed on the test scores of the students. Although the results did not support the hypothesis that a general-to-detailed sequence is superior to a detailed-to-general sequence, an interaction between synthesizer position and the sequence of instructions was found: learning relationships are facilitated when a synthesizer is presented before detailed-to-general instructions or after general-to-detailed instructions. A reference list, five figures, and two data tables are included.—Microfiche 97 cents, paper copy $2.00 plus shipping as document ED 217 859.


This report reviews and evaluates current research and research on the origins, characteristics, and effectiveness of existing approaches to courseware design, and provides suggestions for improving the quality of courseware used in computer based education (CBE). After a brief summary of some of the general problem encountered in courseware design, instructional strategies based on the structure of knowledge, levels of learning, learner differences, and questioning techniques are outlined, and the implications of each type of strategy for courseware development are discussed. A summary of trends indicated by the literature concludes the report, and a 33-item reference list is attached.—Microfiche 97 cents, paper copy $2.00 plus shipping as document ED 217 867.


The purpose of this research project was to develop and evaluate a prototype system-independent Programming Design Guide (PDG) for one of the 13 offline job aids previously developed for use with the Instructional Systems Development (ISD) model. This PDG is intended to provide all of the guidance necessary for a user to implement the job aid on any of a large number of computer systems. The project was divided into three major tasks: (1) the establishment of the content and format for Programming Design Guides; (2) the development of a guide; and (3) the evaluation of the guide. Advice is given for developing Programming Design Guides for other blocks of the ISD model, and ten references are listed.—Microfiche 97 cents, paper copy $3.65 plus shipping as document ED 216 684.

This Programming Design Guide (PDG) was developed to provide computer programmers with all of the guidance necessary for them to implement the offline Job Aid for Selecting Instructional Setting in an inquiry-type version on their computer system. The resulting online program is designed for use by instructional development personnel to assist in the selection of the appropriate instructional setting for each critical task within a Military Occupational Specialty (MOS). As much as the PDG will be used as a guide for programming on a number of computer systems which employ different programming languages, the guide is written in a program design language (PDL) format rather than in any specific computer language. The PDL is a pseudocomputer language which is used to describe the design specification for an interactive computer program.—Microfiche 97 cents, paper copy $10.25 plus shipping as document ED 210 655.


This document provides supplemental information for the user of the online version of the Job Aid for Selecting Instructional Setting; it is not intended to be a stand-alone publication. Included in the guide are sources of information for determining such things as (1) whether there is a considerable amount of theory to be taught with the task under consideration; (2) whether the task must be performed immediately on entry to the job; (3) whether the task is a prerequisite for learning and/or performing other school trained tasks; and (4) whether the task can be learned with very little supervision. Guidelines are also provided for conducting field surveys and organizing panels of experts.—Microfiche 91 cents, paper copy $3.65 plus shipping as document ED 216 666.


This discussion of the uses and methods of evaluation which can be conducted during the development and implementation phases of computer-based instructional (CBI) programs addresses four levels of evaluation: documentation, formative evaluation, assessment of immediate learner effectiveness, and impact evaluation. The components of documentation, which include monitoring project costs, recordkeeping, and personal observation, are outlined; the processes involved in formative evaluation, including internal review and operational testing, are described; the various factors involved in the assessment of immediate learner effectiveness are summarized; and the nature of impact evaluation is discussed.

A set of nine sample evaluation forms relating to the different levels and their associated methods, and a 24-item bibliography are included.—Microfiche 97 cents, paper copy $3.65 plus shipping as document ED 217 870.


The plan described was developed to evaluate the training and cost-effectiveness of the Automated Instructional Management System (AIMS), a computer-based training and resource development system being developed by the Training and Doctrine Command (TRADOC) for possible installation in its schools and training centers. The Functional Description of the proposed AIMS was examined to identify the hardware and software packages intended for development. Stufflebeam's CIPP Model (Context, Input, Process, Product) served as the theoretical framework for the evaluation, and questions were formulated to guide the examination of the AIMS hardware/software, training management, courseware, training effectiveness, cost, and implementation/organization factors. A Milestone Chart was prepared to indicate deadlines for the evaluation activities, and a description of the staff required to implement the evaluation plan was prepared. Draft data collection instruments were devised with the caveat that revisions would be required to tailor them to the specific software packages operational at each AIMS site. A reduced version of this plan has been prepared to evaluate the AIMS test site at the U.S. Army Field Artillery School, Ft. Sill, Oklahoma. Six references are listed, and copies of the data collection instruments are attached.—Microfiche 97 cents, paper copy $10.25 plus shipping as document ED 215 666.


This report describes the use of operations research techniques to determine which courseware packages or what
microcomputer systems best address varied instructional objectives. The MICROPiK model, a highly structured evaluation technique for making such complex instructional decisions, is a multiple alternatives model (MAA) whose overall goal is to formulate an evaluation and decision-making procedure and to model or simulate this evaluation framework as closely as possible while involving the school environment's established needs. An overview of the technical workings of the modeling framework and its performance of the evaluative comparison and final selection of alternative functions is followed by explanation of the primary and secondary goals of the model. Alternatives evaluated by the modeling framework are discussed as well as the criteria necessary to evaluate and compare these alternatives. Additional topics covered include the constraints, execution, results, and general utility of such a model, together with common advantages and potential pitfalls. Microfiche $0.97, paper copy $10.90 plus shipping, as document ED 224 457.


It is argued that, in most existing computer-assisted instruction (CAI), the nature of the instruction would not actually require a computer. Text is presented and read, questions are asked, and then materials are presented or reviewed depending on students' answers. Two problems are found with this approach: computer capabilities for interactive task simulation are underused, and the tacit requirement to present instruction and tests in test form makes learning more difficult for many tasks. These problems may be the reason that little, if any, difference in effectiveness is found between computerized and non-computerized versions of instruction. Recent emphasis in research on mental representations in learning and on work/task simulations for instruction provides a basis for what may be a significant advance in techniques for CAI design. When these ideas are coupled with advances in computer science and technology that make it easier to develop interactive task representations, substantial gains in effectiveness are possible. This paper reviews the problems and selected research and theory, and presents several examples of recent work in CAI implementation that attempt to overcome such problems. Suggestions are derived from this work for the systematic development of design technology for interactive CAI. Microfiche $0.97, paper copy $3.90 plus shipping, as document ED 224 476.


This discussion points out that designing teleconference programs for the physician learner puts unique demands on the teleconferencing medium, as physicians typically expect a 1-hour lecture presentation with high information density. To effectively present the medical content material in the audio medium, strategies which structure and organize the content material are necessary. When high information density programs are teleconferenced, modifications in the format of a typical educational teleconference are necessary. A format which previews, presents, and reviews the content material has proved effective in such programs on the South Dakota Medical Information Exchange (SDMIX) teleconferencing network. This format structures the presentation, making the organization of the content apparent to the physician learner. Visual techniques are used in this organizational pattern to reinforce the structure and facilitate retention of the content presented. The use of a stimulating title slide, a slide of the presenter, program outline, internal visual outlines, frequent changes in the visual elements, summary or transition visuals, color coding of visual and print materials, and humorous visuals have been found effective as means of presenting educational teleconference programs for physicians. Microfiche $0.97, paper copy $2.15 plus shipping, as document ED 224 469.


The faculty training strategy designed by the South Dakota Medical Information Exchange (SDMIX) staff implement

Correspondence Education in the Light of a Number of Contemporary Teaching Models, John A. Baath, Lund, Sweden; Department of Education, University of Lund, 1979. 129pp.

To contribute to a deeper understanding of distance education and its potential development, this 19-chapter book systematically relates correspondence education to seven teaching models: Skinner's behavior control model, Rothkopf's model for written instruction, Ausubel's advance organizer model, Egan's structural communication model, Bruner's discovery learning model, Roger's model for facilitation of learning, and Gagne's general teaching model. Two introductory chapters covering background, aim, and teaching models are followed by in-depth discussions of each of the seven models. Each model is first described, with an explanation of its view of learning and teaching. Then specific ways in which the model could be applied to correspondence education are examined. The final chapter discusses the possibility of applying the models to correspondence education, suggests factors to be considered in choosing and applying models, and describes possible functions of postal two-way communica-

Journal of Instructional Development

A review of the research on correspondence education with emphasis on its two-way communication aspects introduces this report on a series of studies designed to shed light upon problems related to the design and delivery of correspondence instruction. A total of 1,805 adult students enrolled in regular Swedish, Norwegian, and British correspondence courses of a secondary school or vocational training type were randomly assigned to experimental groups with (1) varying "submission density" (equivalent to varying length of study units); (2) varying numbers of assignment questions (omitted questions being replaced by self-check exercises); and (3) traditional postal instruction versus computer-assisted postal instruction. Data were collected through questionnaires, final tests, and a special student register. Results in the third series of experiments showed that the computerized postal instruction was experienced by the students as being more positive than the traditional presentation. Students receiving computer-assisted correspondence tutoring started submitting more assignments than students receiving traditional tutoring by mail. In one of the two experimental courses, they also completed their studies more fully and more rapidly. Examples of postal two-way communication, computer-generated comment letters, and additional tables are appended, and a 13-page reference list is provided. Microfiche $.97, paper copy $2.15 plus postage, as document ED 224 466.


An intelligent videodisc system on which comprehensive instructional development research can be conducted is described. This integrated learning system combines all other existing media except objects, using a videodisc, microcomputer, printer, single monitor, hard disc storage with CPU for random access digitized audio, and headphones. The system components can provide stills or motion with any variation of realism, two audio tracks, any computer-based instruction feature, hard copy, a high resolution color display, and a touch screen. This system will enable researchers to combine design attributes not possible or easily done before, including the combination of realism levels on one visual, complex cueing strategies, and realism combined with a touch screen. A systematic research approach to the identification of effective instructional design techniques is needed and is now possible by varying the videodisc system design features. Dwyer's heart materials would be useful in such research, because they incorporate different levels of objectives and visual abstraction to enhance instruction. His model identifying the types of complexities associated with abstraction level in visuals, modified to account for visual learning processes and combined with the heart materials and the videodisc system, would allow complex investigations. Microfiche $.97, paper copy $2.15 plus shipping, as document ED 223 198.

The above documents may be ordered from the ERIC Document Reproduction Service (EDRS), PO Box 190, Arlington, VA 22210. Please order by ED number, indicate the format desired (microfiche or paper copy), and include payment for the price listed plus shipping. Inquiries about ERIC may be addressed to the ERIC Clearinghouse on Information Resources, School of Education, Syracuse University, Syracuse, NY 13210.

1983, VOL. 6, NO. 3