CHAPTER- II: REVIEW OF LITERATURE

2.0 INTRODUCTION

A review of related literature helps an investigator to define the problem not only with reference to the problem that the investigator wishes to tackle, but also with reference to the topic related to the chosen field. It is an unavoidable step for a researcher as it gives him a deeper insight and understanding of the chosen problem by examining what is already done about the problem the investigator become familiar with various trends of educational phases in the area.

According to Best (1963), “A familiarity with the literature on any problem area helps the students to discover what is already known what others have attempted to find out, what method of attack has been promising or disappointing education which problem remains to solved”.

The present chapter divided into three sections, the first section deals with the studies development of course material, the second section deals with the development of self-instructional material, and the final section concentrated on research studies in distance education with reference to instructional material, quality and credibility.

2.1 STUDIES IN ‘THE DEVELOPMENT OF COURSE MATERIAL’

Schattschneider, Doris. (1998). conducted a study on Development of Course Materials to Integrate Pre calculus Review with the First Course in Calculus. This document describes a two-year project designed to prepare, desk-top publish, and class-test supplemental materials geared specifically toward a slower-paced, one-year Calculus I course that integrates the review of pre calculus topics as they are needed. This course replaces the traditional two-semester Pre calculus-Calculus I sequence for underprepared students. This report includes an executive summary involving project overview, statement of purpose, background and origins, project description, evaluation/project results and conclusions.

Grant, Mary. (1998). Development of a Model Using Information Technology for Support of Rural Aboriginal Students Off-Campus Learning. The study describes the Australian Catholic University's indigenous education program that was developed to serve the education needs of mature aged Aboriginal people, particularly to train teachers and assistants. Topics include course delivery, a model for supported learning using information technology, user characteristics, interactivity, organizational
capacity, and development of course materials

Blackmore, Christine. (1997). Open University Environmental Education and Training. The study describes the approach to environmental education courses at Open University. This includes broad course content, team teaching approach and philosophy of reorienting education towards sustainable development. Course material for open learning includes self-contained study packs as well as course texts, video, audio and computer software. Non formal courses are available to the general public through BBC television programs.

Kennedy, Robert L. (1997). External Validation of an Assessment Class. A survey was conducted to establish the content validity of the material presented during a diagnostic and evaluative procedures course for elementary education majors. Although the content was consistent with information typical of the course, validation through the opinions of practicing teachers could strengthen its validity with the students. An additional benefit is that class members would have testimonials to the value of the class from sources other than their instructor. The 21-question survey was based on the course content outline. Most questions were of the "Do you use..." or "Do you need to know..." variety, but there were several open-ended questions, including one asking for recommendations. Members of 4 assessment classes over 2 years were required to interview from 3 to 10 teachers, depending on the semester, for a total of 333 practitioners from a wide demographic range. Results indicate that the customary topics, including behavioral objectives, Bloom's taxonomy and short-answer test items, are being used. Portfolios and performance testing appear to be well-established, while norm-referenced standardized tests appear to be declining in popularity. Most of the teachers had few content ideas to suggest, possibly indicating that the course content is at least adequate for teachers' needs.

Borzumato, Lawrence P. (1994). The Development of a Learning Assistance Course in Psychology at Ulster County Community College. To address an unacceptable failure and withdrawal rate in an introductory psychology course at Ulster County Community College (UCCC), in New York, a project was undertaken to develop a Learning Assistance in Psychology course to help students with skills and concepts. First, a literature review indicating the need for remediation highlighted the necessity of institutional commitment to a successful program and described characteristics of successful remediation programs. Next, a pilot assistance course was developed and reviewed by the UCCC Skills Center staff and Psychology Department.
The resulting one semester-hour course focused on general study skills, specialized vocabulary, style of quizzes and exams, analysis of learning style, and appropriate note-taking techniques, and was offered to eight students in fall 1993. Five of the eight students achieved a grade of "A" or "A minus," while three withdrew. Informal discussions held with the completing students indicated that more than half believed that participation in the course assisted their understanding of concepts and success. Recommendations from the course validation process included the following: (1) the course should be further analyzed cooperatively among the Skills Center, Psychology Department, and student personnel services; (2) the course should continue to be offered and promoted; and (3) the questionnaire for course participants should be further analyzed and developed.

Black, Harry. et. al. (1993). Evaluation of the Advanced Courses Development Programme. This report presents an evaluation of the Advanced Courses Development Programme (ACDP) that was set up to rationalize the provision of advanced courses in Scottish further education colleges and central institutions. The ACDP is based on students achieving a certain level of competence in a vocational subject defined in terms of outcome and performance criteria and involving student and course assessments. Discussed are background information, course development and validation, the course structure, the benefits and drawbacks of the system. Also examined are local and nationally developed courses and the structure and delivery of the program. The ACDP represents an extensive revision of higher education provision in Scotland, with responsibilities largely devoted to the colleges. The report notes that the phased nature of the implementation has created problems and that guidelines have had to be modified as problems have been encountered and resolved. The report identifies some of these difficulties being faced by staff in colleges as well as the efforts and achievements in addressing them.

Meyer, Luanna H. et. al (1988). A Validation of Program Quality Indicators in Educational Services for Students with Severe Disabilities. A literature search and polling of experts in severe disabilities generated a listing of 123 items representing Program Quality Indicators. Items were then rated by six groups (total N=254) representing such interests as behavior therapy, deaf blindness, mental retardation, state directors of special education, and parents

that is described is divided into subprograms representing different basic numerical concepts—discrimination, recognition, and recall—and presents research design used to test the program's effectiveness with Head Start preschool children. Results indicate that the program was ineffective regardless of length of exposure time.

2.1.1 Summary

In the present section, the studies on development of course material in different disciplines, components in development of course material. Evaluation of course material by using questionnaires, tests, interviews and validation of the course by using internal and external validation were discussed. This section could help the investigator in developing of the course material and validation procedure to see its effectiveness. The studies related to components in development of self-instructional material were discussed in the succeeding section.

2.2 STUDIES IN ‘THE DEVELOPMENT OF SELF - INSTRUCTIONAL MATERIAL’

2.2.1 Teaching Machines

Much of the research regarding Programmed Instruction was based on the use of a teaching machine to implement the instructional event. As Benjamin (1988) noted, “the identification of the earliest teaching machine is dependent on one’s definition of such machines” (p. 703). According to Benjamin’s history, H. Chard filed the first patent for a device to teach reading in 1809. Herbert Akens (a psychologist) patented a device in 1911 that presented material, required a response and indicated whether the response was right or wrong. The contribution of this device, which was a teaching aid rather than an automatic or self-controlling device, was that it was based on psychological research. In 1914, Maria Montessori filed a patent claim for a device to train the sense of touch (Mellan, 1936, as cited in Casas, 1997). Skinner (1958a) and most others (see, for example, Hartley & Davies, 1978) credit Sidney Pressey. Beginning in the 1920s, Pressey designed machines for administering tests. Hartley and Davies (1978) correctly point out that Pressey’s devices were used after the instruction took place, but more important to Skinner, however, was Pressey’s (1926) understanding that such machines could not only test and score they could teach. Moreover, Pressey realized that such machines could help teachers who usually know, even in a small class- room, that they are moving too fast for some students and too slow for others.
2.2.2 Linear Versus Branching Systems

The goal of early developers of programmed instruction was to design the instructional activities to minimize the probability of an incorrect response (Beck, 1959). However, much has been made of the distinction between what some have called Crowder’s (1960) multiple choice branching versus Skinner’s linear-type program (see, for example, Hartley, 1974). Crowder, like Skinner (1954, 1958a) likens his intrinsic system to a private tutor. Although Crowder himself claimed no theoretical roots, his method of intrinsic programming or “branching,” was developed out of his experience as a wartime instructor for the Air Force. Crowder’s method used the errors made by the recruits to send them into a different, remedial path or branch of the programming materials. Although the remediation was not in any way based on any sort of analysis of the error patterns or “procedural bugs” (see, for example Brown & VanLehn, 1980; Orey & Burton, 1992) it may well have been the first use of errors in a tutorial system. Although much has been made about the differences between Skinner and Crowder, it is clear that although the two men worked independently, Skinner was clearly aware of the use of branching and accepted it “if necessary” in 1958 (Skinner, 1958a, p. 974). Crowder began publishing his work a year later in 1959 (Crowder, 1959, 1960, 1964). In a sense they were talking about two very different things. Skinner was writing about education and Crowder was writing from his experience in the teaching complex skills to adults with widely varying backgrounds and abilities. The issue is informative, however. Neither man wanted errors per se. Skinner’s goal was an error rate not to exceed 5 percent (1954). His intention was to maximize success in part in order to maximize (reinforcement) and, at least as important to minimize the aversive consequences of failure. Crowder (1964) would prefer to minimize errors also, although he accepts an 85 percent success rate (15% error rate). Recalling the context of his learner group that ran at least from lectures, such as highly structured sequence of instructional units (frames) with frequent opportunities for the learner to respond via problems, questions, etc. Typically programmed Instruction accompanied by immediate feedback. (p. 3) college graduates to those with an 8 grade education (1964) say, Crowder Lysaught and Williams (1963) suggest that Programmed Instruction maintains the following characteristics. First, it is certainly no one would propose to write materials systematically designed to lead the student into errors and anyone would prefer programs in which no student made an error if this could be achieved without
other undesirable results. We can produce critically effort free programs if we are careful never to assume knowledge that the most poorly prepared student does not have, never to give more information per step than the slowest can absorb, and never to require reasoning beyond the capacities of the dullest. The inevitable result of such programs is that the time of the average and better than average is wasted. (p. 149). In short, Skinner saw errors as a necessary evil motivational and attention getting, but essentially practicing the wrong behavior and receiving aversive consequences for it. Crowder saw errors as unavoidable given the needs of teaching complex skills to students given different backgrounds and whose ability levels varied form “dull” to “better than average.” Crowder’s (1960, 1964) contribution was to try to use the errors that students made to try to find the breakdown in learning or the missing prerequisite skill(s).

2.2.3 Formative Evaluation

Skinner’s (1954, 1958b) early work had indicated the importance of using learner data to make revisions in instructional programs. In a sense, this technology was well established through Tyler’s (1932) discussion of the use of objective based tests to indicate an individual’s performance in terms of the unit, lesson, or course objectives (Dale, 1967). Glaser (1965; Glaser & Klaus, 1962) coined the term criterion-referenced measurement to differentiate between measures concerned with comparing the individual against a criterion score or specific objectives and norm-referenced measurement which ranked the individual’s performance compared to other individuals. What was needed, of course, was to change, atleast in part, the use of such tests from strictly assessing student performance to evaluating program performance. Indeed, Cambre (1981) states that practitioners such as Lumsdaine, May, and Carpenter were describing methodologies for evaluating instructional materials during the Second World War and beyond. What was left was for Cronbach (1963) to discuss the need for two types of evaluation and for Scriven (1967) to label them formative and summative to distinguish between the efforts during development when the product was still relatively fluid or malleable versus the summative or judgmental testing after development is largely over and the materials are more “set.” Markle’s (1967) work became a key reference for the formative and summative evaluation of Programmed Instruction.
2.2.4 Sequencing of Content

Skinner’s (1961) article on teaching machines suggested that the one of the ways that the machine helps with teaching is through the orderly presentation of the program, which in turn is required to be constructed in orderly sequences. Following the selection of an overarching programming paradigm, decisions regarding the sequencing of the content can be made. A general Programmed Instruction program sequence is characterized by an introduction, a diagnostic section, an organizing set/theory section (to help learner focus on primary elements of teaching/testing section), a teaching, testing section, practice section, and finally, a review or summary is presented to reinforce all of the concepts addressed in the specific program (Bullock, 1978). Again, no standard approach exists for the sequencing of content and a variety of models are found in the literature. Lysaught and Williams (1963) describe several techniques, the first of which is the pragmatic approach, or the organization of behavioral objectives into logical sequence. “This order is examined for its internal logic and flow from beginning to end. Often an outline is developed to ensure that all necessary information/steps/components are addressed and that nothing important is omitted” (p. 92). Another common approach to sequencing content was developed by Evans, Glaser, and Homme (1960), and is known as the RULEG system. The RULEG design is based on assumption that material to be programmed consists of rules or examples. So, the rule is presented, followed by examples and opportunities to practice. In some instances, the reverse approach, EGRUL, is used, presenting the learner with a variety of examples and guiding the behavior to comprehend the rule. Mechner (1967) suggests that the target audience should determine which approach is used. If the concept is simple or straightforward, then learners would likely benefit from the RULEG sequence. If the concept is more abstract or complex, then the EGRUL technique would be the better choice in shaping learner behavior. In 1960, Barlow created yet another method for Programmed Instruction design in response to his students’ dislike for the traditional stimulus-response approach, as they felt the technique was too test like. Barlow’s sequencing method was entitled conversational chaining, a reflection of the interconnected nature of the program’s frames. The design requires the learner to complete a response to the given stimulus item, but instead of programmatic feedback about the correctness of that response within the stimulus frame the learner checks his or her accuracy in the following frame. However, the response is not presented
separately, but is integrated within the stimulus of the following frame and is typically programmed Instruction it is initialized so that it is easily identified. As such, the flow of the program is more integrated and capable of eliciting the chain of behavior targeted by the designer. Another well known, but less widely adopted programming method was developed by Gilbert (1962). This approach, called mathetics, is a more complex implementation of reinforcement theory than other sequencing strategies. This technique is also referred to as backwards chaining, since the design is based on beginning with the terminal behavior and working backwards through the process or concept, in stepwise fashion.

2.2.5 Frame Composition

Taber et al. (1965) suggest that a programmed frame could contain the following items: (1) a stimulus which serves to elicit the targeted response (2) a stimulus context to which the occurrence of a desired response is to be learned (3) a response which leads the learner to the terminal behavior and (4) any material necessary to make the frame more readable, understandable, or interesting (p. 90). They also contend that it may not be necessary to include each of these components in every frame. Some frames may contain only information with no opportunity for response, some may be purely directional. One aspect of the stimulus material that is inherent in Programmed Instruction is the inclusion of a prompt. A prompt in Skinner’s view (1957) is a supplementary stimulus, which is added to a program (in a frame or step) that makes it easier to answer correctly. The prompt is incapable of producing a “response by itself, but depends upon atleast some previous learning” (Markle, 1964, p. 36). Skinner proposes two types of prompts, formal and thematic. Formal prompts are helpful in the introduction of new concepts, as learners may have little or no basis for producing their own, unsupported response. A formal prompt typically provided Instruction provides atleast a portion of the targeted response as part of its composition, generating a low-strength response from the learner. Also, the physical arrangement of the frame may serve as a formal prompt type, suggesting to the learner cues about the intended response, such as the number of letters in the response text, underlined words for particular emphasis, the presentation of text to suggest certain patterns, etc. (Taber et al., 1965). Thematic prompts attempt to move the learner toward production and application of the frame’s targeted response in more varied contexts in order to strengthen the learner’s ability to produce the terminal behavior. Taber et al. describe a
variety of design approaches for the creation of thematic prompts. The use of Programmed Instruction structures, grammatical structure, synonyms, antonyms, analogies, rules, and examples are all effective strategies that allow the programmer to create instruction that assists the learner in generating the correct response. The strength of the prompt is another important design consideration and is defined as the likelihood that the learner will be able to produce the targeted response and is influenced by logical and psychological factors related to the design of the frame (Markle, 1964). As new content or concepts are introduced, prompts should be strong to provide enough information so that a correct response can be generated. As low strength concepts are further developed, prompts can be decreased in strength as learners can rely on newly learned knowledge to produce accurate responses. This reduction and gradual elimination of cues is known as fading or vanishing and is another Programmed Instruction related phenomenon popularized by Skinner (1958b).

Another design consideration in the programming of frames is the selection of response type. Taber et al. (1965) describe a variety of response type possibilities and factors related to the basis for selecting from constructed answer, multiple choice, true–false, and labeling, to name a few. Also, another response mode option that has been the subject of instructional research is overt versus covert responding. While Skinner (1968a) believes that active responses are necessary and contribute to acquisition of the terminal behavior, others contend that such forced production may make the learning process seem too laborious (Taber et al.). Research addressing this design issue is described in detail later in this chapter.

2.2.6 Evaluation and Revision

As stated earlier, one of the hallmarks of the Programmed Instruction process is its attention to the evaluation and revision of its products. Skinner (1958a) suggested that a specific advantage of Programmed Instruction is the feedback available to the programmer regarding the program’s effectiveness; feedback available from the learner through trial runs of the product. In fact, many credit Programmed Instruction with the establishment of the first model of instruction that mandates accountability for learning outcomes (Hartley, 1974; Lange, 1967; Rutkaus, 1987). Reiser (2001) indicates that the Programmed Instruction approach is empirical programmed Instruction in nature, as it calls for the collection of data regarding its own effectiveness, therefore allowing for the identification of weaknesses in the program’s design and providing the opportunity
for revision to improve the quality of the program. Markle (1967) presents perhaps the most explicit procedures for three phases of empirical programmed Instruction product evaluation: developmental testing, validation testing, and field-testing. While other authors offer variations on these stages (Lysaught & Williams, 1963; Romiszowski, 1986; Taber et al., 1965), these phases generally represent components of formative and summative evaluation. What factors should one consider when attempting to determine the effectiveness of a program in the production stages? Both Markle (1964) and Lysaught and Williams (1963) indicate that errors in content accuracy, appropriateness, relevance, and writing style are not likely to be uncovered by students in trial situations, and suggest the use of external reviewers such as subject matter experts to assist with initial program editing. Again, Markle (1967) provides the most intricate and rigorous accounts of formative testing, suggesting that once content has been edited and reviewed to address the aforementioned factors, then one-on-one testing with learners in controlled settings should precede field trials involving larger numbers of learners. She insists that only frame-by-frame testing can provide accurate and reliable data not only about error rates, but also information pertaining to communication problems, motivational issues, and learning variables. Some design considerations may cross these three categories, such as the “size-of-step” issue (p. 121), which is both an instructional challenge as well as a motivational factor. Once a program has been produced, many feel that it is the program producer’s obligation to collect data regarding its effectiveness in the field (Glaser, Homme, & Evans, 1959; Lumsdaine, 1965; Markle, 1967). This contention was so compelling that a joint committee was formed from members representing the American Educational Research Association, the American Psychological Association, and the Department of Audiovisual Instruction (a division of the National Education Association). The report created by this Joint Committee on Programmed Instruction and Teaching Machines (1966) offers guidance to a variety of stakeholders regarding the evaluation of program effectiveness, including programmatic effectiveness data that prospective purchasers should seek, as well as guidelines for program producers and reviewers in their production of reports for the consumer. While the committee expresses the value inherent in one-on-one and small group testing, they place stronger emphasis on the provision of data from larger groups of students and repeated testing across groups to demonstrate the program’s reliability and validity in effecting its intended outcomes. In his description of considerations for program assessment, Lumsdaine (1965) is careful to point out the
need to distinguish between the validation of a specific program and the validation of Programmed Instruction as an instructional method, a distinction that has continued through present day evaluation concerns (Lockee, Moore, & Burton, 2001). Although evaluation and research may share common data collection approaches, the intentions of each are different, the former being the generation of product-specific information and the later being concerned with the creation of generally applicable results, or principles for instruction (Lumsdaine, 1965).

### 2.2.7 Pressey’s teaching Machines

Pressey’s Machines, Pressey’s self-instruction devices were developed to provide students with immediate feedback of results on knowledge after reading and listening to a lecture. Most of the research on Pressey’s devices dealt with implementation and use of the results in order to develop a specific type of information to help the instructor change content and approach. Stolurow (1961) raised a question early on: when a programmed machine is used in conjunction with other means of instruction, which would be the cause of any effect? He felt it would be important to be able to judge how effective the programmed devices would be when used alone versus when they were used in conjunction with other types of instruction.

There was less concern about the problems of programming and sequencing in these machines (Stolurow, 1961). An example of research in this category was Peterson (1931) who evaluated Pressey’s concepts with matched participants who were given objective pre and post-tests. The experimental group was given cards for self-checking their responses while the control group received no knowledge of feedback. In another version the participants were given a final test that was not the same as the post tests. In both situations the experimental group with knowledge of results scored higher than the control group. Little (1934) compared results from groups either using a testing machine, a drill machine, or neither (control group). Both experimental groups scored significantly higher than the control group. The group using the drill machine moved further ahead than did the test machine group. The group using the drill machine moved further ahead than did the test machine group. Cassidy (1950), a student of Pressey) in a series of studies on the effectiveness of the punch board reported that the immediate knowledge of results from this device provided significant increments in the learning of content. Pressey (1950) conducted a series of studies used punch board concepts at the Ohio State University designed to test whether punchboard teaching machines could
produce better learning performance by providing immediate knowledge of results and whether these beneficial effects are limited to a particular subject (Stolurow, 1961, p. 105). This series of studies lead to the following conclusions by Pressey and his associates as reported by Stolurow.

1. The use of the punchboard device was an easy way of facilitating learning by combining feedback, test taking, and scoring.
2. Test taking programs could be transformed to self directed instruction programs.
3. When punch boards were used systematically to provide self instruction, content learning was improved.
4. Automatic scoring and self instruction could be achieved by the use of the punchboard.
5. The technique of providing learners with immediate knowledge of results via the punchboard could be used successfully in a variety of subjects. (1961).

Stephens (1960) found that using a Drum Tutor (a device used with informational material and multiple-choice questions and designed that students could not progress until the correct answer was made) helped a low-ability experimental group to score higher on tests than a higher ability group. This study confirmed Pressey’s earlier findings that “errors were eliminated more rapidly programmed Instruction with meaningful material and found that students learned more efficiently when they could correct errors immediately” (Smith & Smith, 1966, p. 249). These data also suggested that immediate knowledge of results made available early within the learning situation are more effective than after or later in the process (Stolurow, 1961). Severin (1960), another student of Pressey, used a punchboard testing procedure to compare the achievement of a learners forced to make overt responses versus those who were not required to make overt responses. No differences were reported. He concluded on short or easy tasks the automated overt devices which were of little value. In an electrified version of the Pressey punchboard system, Freeman (1959) analyzed learner performance in a class of students who received reinforcement for a portion of the class and no reinforcement for another portion of time. He found no significant effects related to achievement; however, he indicated that in this study there were problems in the research design, including insufficient amount of reinforced opportunity, that test items were not identical to reinforced ones, and there was little attempt to program or structure the reinforced test materials (items). Freeman also noted that rapid programmed Instruction gains in learning might not relate to better retention.
Holland (1959), in two studies on college students studying psychology using machine instruction, required one group of students to space their practice versus another group of students who had to mass their practice. He reported no significant differences as a result of practice techniques.

Stolurow (1961) suggested that studies on Pressey’s machines, as a way of providing learners with immediate knowledge of results indicated that these machines could produce significant increments in learning, that learning by this method was not limited to particular subject areas and that the approach could be used with various types of learners. The effectiveness of having knowledge of results made available by these machines depended a great deal upon how systematic the material was programmed, the type of test to determine retention, and the amount of reinforced practice. Smith and Smith (1966) and Stolurow (1961) indicated that, based upon reviews of Pressey’s earlier experiences, that there are positive outcomes of machine-based testing of programmed material. However, they also contended that the programmed machines may be more useful when used in connection with other teaching techniques. Pressey (1960) himself, states, “Certainly the subject matter for automation must be selected and organized on sound basis. But the full potentialities of machines are now only beginning to be realized” (pp. 504–505). In reference to the effectiveness of programs on machine, Stolurow (1961) concluded that they are effective in teaching verbal and symbolic skills and for teaching manipulative skills. Please note that there is a great overlap of the research on programmed machines and materials and of other approaches and variations. Additional programmed machine research is re-viewed later in this section to illustrate points, concerns, and applications of other programming variables and research.

2.2.8 Military Knowledge Trainers

A major design and development effort in the use of automated self instruction machines was conducted by the U.S. Air Force, Office of Naval Research and by the Department of Defense during and after World War II. These development projects incorporate the concepts of Pressey’s punchboard device in the forms of the Subject Matter Trainer (SMT), the Multipurpose Instructional Problem Storage Device, the Tab-Item, and Optimal Sequence Trainer (OST), and the Trainer-Tester (see Briggs, 1960 for a description of these devices). These automated self instructional devices were designed to teach and test proficiency of military personnel. The Subject Matter
Trainer (SMT) was modified to include several prompting, practice, and testing modes (Briggs, 1956, 1958). The emphasis of the SMT was to teach military personnel technical skills and content (Smith & Smith, 1966). Bryan and Schuster (1959) in an experiment found the use of the OST (which allowed immediate knowledge following a specific response) to be superior to regular instruction in a troubleshooting exam. In an experimental evaluation of the Trainer-Tester and a military version of Pressey’s punchboard, both devices were found to be superior to the use equipment mock-ups and of actual equipment for training Navy personnel in electronic troubleshooting (Cantor & Brown, 1956; Dowell, 1955). Briggs and Bernard (1956) reported that an experimental group using the SMT, study guides, and oral and written exams outperformed the control group who used only the study guides and quizzes on a performance exam. However, the two groups were not significantly different on written tests. Both of these studies were related to the extent to which instruction provided by these machines was generalizable or transferable. With respect to the effectiveness of these versions of teaching machines, these studies indicated that these programmed machines (SMT) can “be effective both for teaching verbal, symbolic skills which mediate performance and for teaching overt manipulative performance” (Stolurow, 1961, p. 115). Not all studies, however, reported superior results for the Subject Matter Trainer. He pointed out that these devices, which used military content and subjects generally, showed a consistent pattern of rapid programmed Instruction learning at various ability levels and content and suggested that knowledge of results (if designed systematically) was likely to have valuable learning benefits.

2.2.9 Skinner’s teaching Machines

The research studies on Pressey’s punchboard devices, and their military versions (e.g., SMT, OST, etc.), which incorporated many features of self-instruction and supported the concept that knowledge of results would likely have beneficial educational applications. However, the real impetus to self-instruction via machine and programmed instruction came from the theories and work of B.F. Skinner (e.g., 1954, 1958, and 1961). Skinner’s major focus was stating that self-instruction via programmed means should be in the context of reinforcement theory. He felt that Pressey’s work was concerned “primarily with testing rather than learning and suggested that the important ideas about teaching machines and programmed instruction were derived from his analysis of operant conditioning” (Smith & Smith,
1966, p. 251). (See descriptions of these devices earlier in this chapter.) Skinner described his devices similar to Pressey's descriptions, including the importance of immediate knowledge of results. The major differences were that Pressey used a multiple-choice format and Skinner insisted upon constructed responses, because he felt they offered less chance for submitting wrong answers. Skinner's machines were designed to illicit overt responses. However, his design was modified several times over the years allowing more information to be presented and ultimately sacrificed somewhat the feature of immediate correction of errors. Skinner was most concerned about how the materials were programmed to include such concepts as overt response, size of steps, etc. As a result, much of the research was conducted on these programming components (concepts). These programming features included presenting a specific sequence of material in a linear, one-at-a-time fashion, requiring an overt response and providing immediate feedback to the response (Porter, 1958). Research on these components will be discussed later in this chapter. Much of the literature on Skinner's machines was in the form of descriptions of how these machines were used and how they worked (e.g., Holland, 1959; Meyer, 1959).

2.2.10 Crowder's Intrinsic Programming

Crowder (1960) modified the Subject Matter Trainer to not only accommodate multiple choice questions, but to include his concept of branching programming in “which the sequence of items depends upon the response made by the student. Correct answers may lead to the drop programmed Instructioning of certain items, or incorrect answers may bring on additional remedial material” (Smith & Smith, 1966, p. 273). Crowder's theories, like Skinner's were not machine specific. Much of the research was based around the various programmed aspects noted above. These programming aspects (variations) espoused by Crowder (1959, 1960) (e.g., large blocks of information, branching based upon response, etc.) will be also reviewed later in this chapter.

2.2.11 Ability, Individual Differences and achievement

Glaser, Homme, and Evans (1959), suggested that individual differences of students could be important factor based upon previous research, which might affect program efficiency. Several questions arise under these assumptions: (1) Does student ability (or lack of) correlate with performance in a programmed environment? (2) Does performance in a programmed environment correlate with conventional instructional
methods and settings? Again, there appears to be no consensus in the results or the recommendations of the research.

Porter (1959) and Ferster and Sapon (1958) reported in separate studies that there was little or no correlation between ability level and achievement on programmed materials. Detambel and Stolurow (1956) found no relationship between language ability and quantitative subtests of ACE scores (American Council on Education Psychological Examination for College) and performance on a programmed task. Keisler (1959) matched two groups on intelligence, reading ability, and pre test scores, with the experimental group using a programmed lesson; the control group received no instruction. All but one of the experimental subjects scored higher after using the programmed materials. Two groups of Air Force Programmed Instruction lots were matched according to duties, type of aircraft, and “other” factors, with one group having voluntary access to a programmed self-tutoring game on a Navy Automatic Rater device. After two months the experimental group with voluntary access to the programmed materials showed significant improvement on items available with the game. The control group did not show significant improvement. However, there was no difference between the groups on items not included in the programmed materials. It was concluded that a self-instructional device would promote learning even in a voluntarily used game by matched subjects (Hatch, 1959).

Dallos (1976) in a study to determine the effects of anxiety and intelligence in learning from programmed learning found an interesting interaction on difficult programs. He reported that a high state of anxiety facilitated learning from the higher intelligence students and inhibited learning for low intelligence students.

Carr (1959) hypothesized that effective self instructional devices would negate differences in achievement of students of differing aptitudes. Studies by Porter (1959), and Irion and Briggs (1957), appeared to support this hypothesis as they reported in separate studies little correlation between intelligence and retention after using programmed devices. Carr (1959) suggested that the lack of relationship between achievement and intelligence and/or aptitude is because programmed instruction renders “learners more homogeneous with respect to achievement scores” (p. 561). Studies by Homme and Glaser (1959), and Evans, Glaser, and Homme (1959) tended to also support Carr’s contention, while Keisler (1959) found students using machine instruction were more variable on achievement scores than the control group not using the programmed machines. Carr (1959) called for more study to determine
the relationship between achievement and normal predictors with the use of programmed instruction.

2.2.12 User Attitude

Knowlton and Hawes (1962) noted, “that the pull of the future has always been slowed by the drag of the past” (p. 147). But, as there is a resistance to new technology, what proves valuable is thus too accepted. This statement appears to sum up the attitude toward programmed instruction in that perception of problems is due to lack of relevant information by the programmers and researchers.

Smith and Smith (1966) reported that the general reaction of learners towards programmed instruction at all levels including adult learners was very positive. This view was borne out by a number of studies gauging attitudes of learners toward programmed self-instruction. Stolurow (1963), in a study with retarded children using programmed machines to learn mathematics, found that these students, while apprehensive at first, later became engrossed and indicated they preferred using the machines rather than having traditional instruction. However, Porter (1959), in his earlier noted study, reported that there was no relationship among the gender of the student, the level of satisfaction with the programmed method, and achievement level. Students in a high school study revealed a view that was balanced between the use of programmed programs and conventional instruction (First Reports on Roanoke Math Materials, 1961). Eigen (1963) also reported a significant difference between attitudes of the use of programmed materials and other instruction of 72 male high school students in favour of the programmed instruction. Nelson (1967) found positive attitudes in student perceptions of programmed instruction in teaching music. Likewise, several studies on attitude were conducted in college classrooms. Engleman (1963) compared attitudes of 167 students using programmed and conventional instruction (lectures, labs, etc.) and reported that 28 percent indicated programmed materials were absolutely essential, 36 percent felt they were useful 90 percent of the time, 21 percent considered programmed materials useful 50 percent of the time, and 14 percent indicated that programmed materials were helpful only occasionally or not at all. Cadets at the Air Force Academy showed moderate enthusiasm as 80 percent indicated enjoyment in the programmed course, however, 60 percent preferred it to conventional teaching and suggested they learned with less effort (Smith, 1962). Several opinion programmed Instruction studies were conducted in three colleges (Harvard, State College at Genesco, and Central
Washington University) comparing attitudes of students using a programmed text, The analysis of behavior, (Holland & Skinner, 1961) and a textbook entitled A textbook of psychology (Hebb, 1958). The attitudes were overwhelming positive toward the programmed text (Naumann, 1962; Vanatta, 1961). Skinner and Holland (1960) reported that 78 percent of the students “felt they learned more from the machine than from the text” (p. 169). Banta (1963) reviewed similar attitude measures at Oberlin, University of Wisconsin and Harvard and results were somewhat less favorable than the above study, but the Harvard students’ attitude scores were similarly positive. Smith and Smith (1966) speculate that because the materials were developed at Harvard, there may have been a tendency to reflect their teachers’ “enthusiasm and reacted in the expected manner” (p. 302). Roth (1963) also reported results of another college graduate students’ opinion programmed Instruction of the same Holland and Skinner text. All students liked it in the beginning, but only five did at the end of the study. Several objections noted that the program was “tedious,” “repetitive,” “mechanized,” “non-thought provoking,” and “anti-insightful” (Roth, 1963, p. 279–280). In a business setting at IBM, Hughes and McNamara (1961) reported that 87 percent of trainees liked programmed materials better than traditional instruction. Tobias (1969a, 1969b) provided evidence that teacher and user preferences for traditional devices are negatively related to achievement in programmed instruction. There have been a variety of studies dealing with student attitude toward various aspects of the programming variables. Jones and Sawyer (1949), in a study comparing attitudes of students using a programmed machine which provided self scoring and immediate knowledge of results versus a conventional paper answer sheet found 83 percent preferred the machine program over the paper answer sheet. Two studies (Eigen, 1963; Hough & Revsin, 1963) reported conflicting results on positive attitudes toward programmed machine and programmed texts. In a study concerning anxiety and intelligence when using difficult programmed instruction, Dallos (1974) found that participants with high anxiety, but lower intelligence had unfavourable view of the programmed instruction while the high intelligent, high anxiety participants had more favorable opinions programmed Instruction of the program. Studies on attitude and learning effectiveness of programmed instruction have indicated that positive or negative attitudes toward programmed materials have little or no predictive value in determining learning effectiveness of these programs (Eigen, 1963; Hough & Revsin, 1963; Roe, Massey, Weltman, & Leeds, 1960; Smith & Smith, 1966). Smith and Smith (1966) indicated
that these findings were not surprising because of other studies on general behavior have shown similar results (e.g., Brayfield & Crockett, 1955). “The apparent fact is that general attitude measures predict neither learning nor performance in a particular situation” (Smith & Smith, 1966, p. 304).

2.2.13 Conventional Instruction (Comparison Studies)

Much of the research on programmed machine and programmed instruction involved comparing programs to conventional or traditional instruction (whatever that was or is). This comparison technique was flawed from the beginning, but the results using this technique were used by many as proof the program was successful or was a failure, or was it just as good as the other form of instruction (incorrectly interpreting the no significant difference result). Anytime one method of instruction is compared with another, several issues need to be kept in mind. Sometimes the comparisons are made between small groups with limited content and for relatively short time. Secondly, the novelty may effect operates in many cases generally supporting the new technique, e.g., programmed instruction. Thirdly, there are many uncontrolled factors operating all at once and any of these may affect the results of the study (Smith & Smith, 1966). This noted, in a review of 15 studies comparing programmed and conventional instruction, Silberman (1962) reported that nine favoured programmed instruction and six indicated no significant difference in the two approaches. All 15 studies reported that the programmed approach took less time. Several studies reported that when specific content was taught using programmed methods, time was saved with no decrease in achievement. All reported that instruction time was saved or the program instruction completed requirements in less time than a conventional group (Hosmer & Nolan, 1962; Smith, 1962; Uttal, 1962; Wendt & Rust, 1962). In a study to compare a traditional instruction to a programmed method of teaching spelling in the third grade, the programmed group gained significantly better grade-equivalent scores than the control group by the end of the year (Edgerton & Twombly, 1962). Hough (1962) compared machine programs to conventional instruction in a college psychology course where time was an additional factor. When quizzes were not announced the machine instructed group scored significantly higher, but when quizzes were announced, there was no significant difference. Since the conventional group could study at home, whereas the machine group could not, the additional time available to the conventional group was a factor in these results.
Hartley (1966, 1972) reviewed 112 studies that compared programmed instruction (any variety) and conventional instruction. He concluded that there is evidence that programmed instruction is as good, or more effective than conventional instruction. In addition, Hamilton and Heinkel (1967) concurred in Harley’s findings, which found in 11 of 12 studies that compared an instructor with a programmed lesson, an instructor alone, or a program alone, that an instructor with a program was the more effective choice. Hartley (1978) states “the results allow one to make the generalizations that many programs teach as successfully as many teachers and sometimes that they do this in less time” (p. 68). Falconer (1959) believed that it is an advantage for deaf children to use teaching machines where they traditionally require a large amount of individual instruction. He suggested that his data indicated that a teaching machine might be as effective as a teacher who had to spread his/her time over many students individually. Day (1959) compared a group using a Crowder style programmed book with that of conventional instruction. The experimental group that used the programmed book scored 20 percent higher and made one-fourth the wrong answers than the conventional instruction group over a half semester course. Goldstein and Gotkin (1962) reviewed eight experimental studies, which compared programmed text to programmed machines. Both versions were linear in nature. Goldstein and Gotkin reported no significant differences on several factors; post test scores, time, and attitude across both presentation modes. (Four studies indicated the programmed texts used significantly less time than the machine version, however.) Other studies have shown no significant difference between automated instruction and traditionally taught classes or were equally effective modes of instruction (Goldberg, Dawson, & Barrett, 1964; Oakes, 1960; Tsai & Pohl, 1978). Similar no significant difference results were reported in studies with learning disabled students (e.g., Blackman & Capobianco, 1965; McDermott & Watkins, 1983; Price, 1963). Porter (1959) did report results showing that second and sixth graders progressed further in spelling achievement with programmed materials in less time than in a conventional classroom setting.

Silberman (1962) reviewed eight comparative studies to determine how best to present material in a self-instruction program, e.g., small step, prompting, overt response, branching, or repetition. He reported that there was no clear pattern of success and these cases showed that some treatments favoured one method or another while other treatments favoured the time-on-task factor. There were no significant differences across the programmed modes. Eighth grade students of high ability were
put into three groups, one used a linear program, other used a branching program, and the third was used as a control group (conventional instruction). Time available was constant across all groups. In a result unusual for this type of study, Dessart (1962) reported that the control group did significantly better than the experimental group using the branching approach. There was no significant difference between the conventional group and the linear group or between the linear and branching groups.

Stolurow (1963) studied the effect of programs teaching learning disabled children reading, vocabulary, and comprehension. Although, the results favoured the programmed version over a traditional method, Stolurow recommended altering programs with conventional instruction. His recommendation was similar to others, which suggested a variety of methods may be more effective than only using one. Klaus (1961) reported on a comparison study dealing with 15 high school physics classes. Some classes had programmed materials available but not for mandatory use. The class having access to the programs had a substantial gain in criterion scores compared to the class without these materials available. After reviewing several studies, Alter and Silverman (1962) reported there were no significant differences in learning from the use of programmed materials or conventional texts. McNeil and Keisler (1962), Giese and Stockdale (1966), Alexander (1970), and Univin (1966) in studies comparing the two versions (programmed and conventional texts) also found the similar results of no significance across methods. However, in a number of studies using primarily retarded learners, the reported results of these comparison studies found the conventional instruction to be superior (Berthold & Sachs, 1974; McKeown, 1965; Richmond, 1983; Russo, Koegel, & Lovaas, 1978; Weinstock, Shelton, & Pulley, 1973). However, the programmed devices (particularly linear ones) have the advantage over teachers in a conventional setting who, in some cases, inadvertently skip over small ideas or points, which may need to be present for understanding. Some feel these programmed devices could solve this concern (Stolurow, 1961). When program machines were studied as the sole source of instruction, Stolurow (1961) indicated in his review that both children and adults benefited from a programmed device. He stated, “These devices not only tend to produce performance which is freer of error than conventional methods of instruction, but also reduce the amount of instruction time required” (p. 135–136).
2.2.14 Programmed Variables (Essential Components)

During the early development of programmed instruction devices and materials many ideas were expressed on how best to present information, some based in theory (e.g., Skinner’s work), others based on intuition, but little on actual research. Reviews of existing literature (e.g., Silberman, 1962) yielded no clear pattern of what programming criteria was effective in improving achievement. However, as time passed more studies and analysis of programming variables were conducted. Program or programming variables are components that are essentially general in nature and can be associated with all types of programs. For an example, these variables can deal with theoretical issues such as the effect overt versus covert responses, the impact of prompting or no prompting, size of steps, error rate, or the confirmation of results. Other issues indirectly related to the programming variables include user attitudes toward programs the mode of presentation (e.g., linear and branching) and program effectiveness. Illustrative results are provided from representative research studies.

2.2.15 Mode of Presentation

Various studies have been conducted comparing linear to branching programs, both in terms amount of learning and time saved in instruction. Coulson and Silberman (1960), and Roe (1962) found no significant differences in test scores between the two versions, but both found significant differences in terms of time taken to learn favouring branching programs. However, Roe (1962) did find that forward branching and linear programs were significantly faster (in terms of time saved) than backward branching. Mixed results were found in other studies, for example, Silberman, Melaragno, Coulson, and Estavan (1961) found no significant difference between the versions of presentation on achievement, but in the following study, Coulson, Estavan, Melaragno, and Silberman (1962) found that the branching mode was superior to a linear presentation.

Holland (1965), Leith, (1966), and Anderson (1967) reported no significant difference in learning between linear and branching programs when compared, and indicated this was generally the case with older or intelligent learners, “younger children using linear programs were more likely to receive higher test scores, although often these still took longer to complete than did branching ones” (Hartley, 1974, p. 284).
2.2.16  Overt Versus Covert Responses

One of Skinner’s principles of programmed instruction is the necessity of overt responses. It appeared to be an important research concern to determine when it is advantageous to require overt or allow covert responses that could affect learning achievement. Are covert responses as effective as overt ones? This question has been a popular research topic programmed Instruction. Overt responses require the student to do something (e.g., writing or speaking an answer, while covert requires thinking about or reading the material). Skinner’s (1958) theory requires that a response should be overt (public) because if not overt, responses often ceased (Holland, 1965). Holland (1965) suggested that covert responses are not necessarily theoretical but also practical, because all aspects (in Skinner’s view) of a program necessitate getting the correct answer. “Therefore, a measure of a program by not answering at all circumvents the characteristics which make it a program” (p. 93). Holland (1965) continued, indicating that several conditions must be met to determine the difference between overt and covert responses, namely, (1) program design must allow the student to answer correctly, and (2) the correct answer can only be attained after the appropriate steps in the program have been completed. Other researchers over the years have accepted this concept as important (e.g., Tiemann & Markle, 1990). In reviews of research by Lumsdaine (1960, 1961), Feldhusen (1963), and Silberman (1962), all reported some mixed results, but the overall finding was that there was no difference in achievement between the overt or covert response groups. Results of several studies suggest that the use of overt responses was supported under some conditions (e.g., Briggs, Goldbeck, Campbell, & Nichols, 1962; Williams, 1963; Wittrock, 1963). Holland (1965) reported that when answers on a test are not contingent on important content, overt responding might not be effective. Otherwise, studies indicated a test advantage for students using overt responses. Goldbeck and Campbell (1962) found that the advantages of each type of response may vary with the difficulty of content. Additionally, several studies showed that overt responding in programmed instruction was beneficial over covert responses (Daniel & Murdock, 1968; Karis, Kent, & Gilbert, 1970; Krumboltz & Weisman, 1962; Tudor, 1995; Tudor & Bostow, 1991; Wittrock, 1963). Miller and Malott (1997) in a review of the literature on effectiveness of overt responses versus non overt responses concluded that there was little benefit in requiring overt responses.
when additional learning-based incentives are present, but in situations where no incentives are present overt learning should improve learning.

A large number of other researchers found no significant difference between the effectiveness of programmed materials requiring overt responses and those using covert responses (Alter & Silberman, 1962; Csanyi, Glaser, & Reynolds, 1962; Daniel & Murdock, 1968; Goldbeck & Campbell, 1962; Goldbeck, Camp- bell, & Llewellyn, 1960; Hartman, Morrison, & Carlson, 1963; Kormandy & VanAtta, 1962; Lambert, Miller, & Wiley, 1962; Roe, 1960; Stolurow & Walker, 1962; Tobias, 1969a, 1969b, 1973; Tobais & Weiner, 1963). Shimamune (1992) and Vunovick (1995) found no significant difference between overt construction and discrimination responses and covert responses. However, in these studies extra credit (incentives) was given for test performance. Miller and Malott (1997) replicated Tudor’s (1995) study and found that the no-incentives overt group produced greater improvement than did the covert responding group. This was also true for the incentive overt responding group as well. Their results did not support earlier studies (noted above) and concluded that overt responding was “robust enough phenomenon to occur even when an incentive is provided” (p. 500).

Evans et al. (1959) required two groups to use machine instruction except one group was required to answer overtly, the other group were required not to answer items overtly. They reported no significant difference in the approach, but the non overt answering group took less time than the overt group. While the research reported primarily no significant difference between learners who wrote answers and thought about answers, Holland (1965), Leith (1966), and Anderson (1967) felt that there were situations in which overt answers were superior to covert answers. Hartley (1974) summarized these situations: (1) when young children were involved, (2) when materials were difficult or complex, (3) when programs were lengthy, and (4) when specific terminology was being taught. There is, however, evidence according to Glaser and Resnick (1972), and Prosser (1974) the mere questioning is important to learning, regardless of covert or overt response situations.

2.2.17 Prompting

Holland (1965) indicated that in a study of paired associates, prompting was defined as a response given prior to an opportunity to have an overt response, whereas when confirming the response item is given after the overt response. Several studies
dealt with the advantages of prompting versus non-prompting in a program sequence. Cook and Sttpzer programmed Instruction (1960) and Cook (1961) reported a no significant difference between the two versions, and also indicated that overt responses were not necessary for better achievement. Angell and Lumsdaine (1961) concluded from the review several studies that programs should include both prompted and non prompted components. Stolurow, Hasterok, and Ferrier (1960) and Stolurow, Peters, and Steinberg (1960) in preliminary results of a study reported the effectiveness of prompting and confirmation in teaching sight vocabulary to mentally retarded children. In an experiment comparing a partial degree of prompting (prompting on 3/4 of the trials) to a complete prompting (prompting on every trial) version, Angell and Lumsdaine (1961) found learning was significantly more efficient under the partial prompting condition and supported the results of Cook (1958) and Cook and Sttpzer programmed Instruction (1960).

2.2.18 Confirmation

There appears to be some controversy over the concept or interpretation of feedback, reinforcement, and confirmation. Skinner (1959) interpreted confirmation as a positive reinforce in the operant conditioning model (Smith & Smith, 1966). Others have objected to this view suggesting that getting a student to perform a desired function for the first time is not addressed (Snygg, 1962). Lumsdaine (1962) suggested that program developers should be most interested in the manipulation of prompting cues, not manipulation of reward schedules. Smith and Smith (1966) indicated that in an operant conditioning situation the response and the reinforcement are constant while in programmed instruction the situations are continually changing. Several studies compared programs with confirmation (after an overt answer, the correct answer is presented) to programs with no confirmation available. No significant difference was found in scores as a function of confirmation (Feldhusen & Birt, 1962; Holland, 1960; Hough & Revis, 1963; Lewis & Whitwell, 1971; McDonald & Allen, 1962; Moore & Smith, 1961, 1962; Widlake, 1964). However, Meyer (1960), Angell (1949), and Kaess and Zeaman (1960) found significant advantages in answer confirmation. Suppes and Ginsberg (1962) found an overt correction after confirmation to be also effective. Krumboltz and Weisman (1962) in comparing continuous versus non continuous confirmation, reported neither had an effect on the test scores. Repetition and review have been built into many programs. Some programs were designed to drop a question
when it had been correctly answered. Because it was technically easier in 1960s to drop out a question after only one correct response rather than after additional responses, many programs were designed this way. However, Rothkopf (1960) did try to determine if there was any advantage to dropping programmed Instruction questions out after two correct responses or any advantage to a version where none of the questions were dropped. He reported that the two methods were equally effective. Scharf (1961) and Krumboltz and Weisman (1962) investigated several schedules of confirmation and found no significant difference. However, Holland (1965) claimed even in the absence of significant results, that there was “enough suggestion of small differences so that the importance of confirmation cannot be discounted” (p. 91). Jensen (1949), Freeman (1959), and Briggs (1949) all reported that when there is a frequent, deliberate, and systematic effort to integrate the use of knowledge of results, learning shows a cumulative effect in a significant manner. Hartley (1974) in his review and summary of programmed learning research on learner knowledge of results argued that immediate knowledge affected some learners more than others. In experiments “with low ability learners and with programs with higher error rates, immediate knowledge of results was found to be helpful” (Holland, 1965; Anderson, 1967; Annett, 1969, as cited in Hartley, 1974, p. 284). Although reinforcement, feedback, and confirmation are central issues to programmed instruction research, this area of research is incomplete and additional information concerning variables such as amount, schedule, and delay of reinforcement was missing. There appears to be no research that explains the problem of why confirmations are not always needed or why programs exhibiting the “pall effect” (boredom induced by the program) could promote learning (Rigney & Fry, 1961, p. 22).

2.2.19 Sequence

The basic structure of programmed machines and materials is a systematic progression of behavioural steps, which takes the student through complex subject matter with the intention of knowledge acquisition. One of Skinner’s major tenants was the “construction of carefully arranged sequences of contingencies leading to the terminal performance which are the object of education” (Skinner, 1953, p. 169). This sequence of information and progressions in terms of “both stimulus materials displayed to the student and the way in which he interacts with and responds to them” are a fundamental issue of programmed learning research (Taber et al., 1965, p. 167).
Gavurin and Donahue (1960) compared a sequenced order of a program with a scrambled order version as to the number of repetitions required for an errorless trial and on the number of errors to reach criterion. For both measures the sequenced order was significantly better. Hickey and Newton (1964) also found a significant difference in favour of original sequence to another unordered one. Hartley (1974) indicated that this suggested that the “analysis of structure must be very sophisticated indeed if it is to reveal useful differences in sequencing procedures” (p. 283). Roe, Case, and Roe (1962) found no significant difference post-test scores on a scrambled ordered versus a sequenced ordered program on statistics. However, using a longer form of the same program, Roe (1962) found significant advantages for the ordered sequences, on the number of student errors on the program and amount of time needed to complete the program. Several research studies comparing ordered program sequences with non logical or random sequences have not supported Skinner’s principle of ordered sequences (Duncan, 1971; Hamilton, 1964; Hartley & Woods, 1968; Miller, 1965; Neider-meyer, Browen, & Sulzen, 1968; Wager & Broaderick, 1974). However, Wodkte, Brown, Sands, and Fredericks (1968) found some evidence that the use of logical sequences for the lower ability learner was positive. Miller’s (1969) study indicated that logical sequence appears to be the best in terms of overall effectiveness and efficiency. He felt it would be of value, however, to identify which levels of sequencing would be the most effective. In a review of several studies on logical sequencing, Hartley (1974) indicated that learners could tolerate “quite considerable distortions from the original sequence and that the test results obtained are not markedly different from those obtained with the original program’s so called logical sequence” (p. 282). He stressed that these studies were conducted on short programs, however.

### 2.2.20 Size of Step

Size of step generally refers to the level of difficulty of the content or concepts provided in a frame. In addition, step size can mean, (1) amount of materials, for example, number of words in a frame, (2) difficulty as in error rate, and (3) number of items present (Holland, 1965). Thus, research in this category varies by “increasing or decreasing the number of frames to cover a given unit of instruction” (Smith & Smith, 1966, p. 311). Using a programmed textbook with four levels of steps (from 30 to 68 items), four groups of students completed the same sequence of instruction, each group
with a different number of steps. Evans et al. (1959) reported in that the group using smaller steps produced significantly fewer errors on both immediate and delayed tests. Likewise, Gropper (1966) found that larger the step size, the more errors were committed during practice. This finding was significant for lower ability students.

Smith and Moore (1962) reported in a study in which step size (step difficulty) and Programmed Instruction cues were varied in a spelling program, that no significant difference was found on achievement related to step size, but the larger step program took less time. Smith and Smith (1966) opened programmed Instruction, “very small steps and over cueing may produce disinterest” (p. 311). Balson (1971) also suggested that programmers could “increase the amount of behavioural change required of each frame” and thus increase the error rate, but not decrease achievement levels and also have a significant saving of time in learning (p. 205). Brewer and Tomlinson (1981) reported that except for brighter students, time spent on programmed instruction is not related to either improvement in immediate or delayed performance. Shay (1961) studied the relationship of intelligence (ability level) to step size. He reported relationship and indicted that the small steps were more effective (producing higher scores) at all ability levels.

Rigney and Fry (1961) summarized various studies and indicated that programs using very small (many components to a concept) could introduce a “pall effect” (Rigney & Fry, 1961, p. 22) in which boredom was inducted by the material, particularly with brighter students. These results were later supported by Briggs et al. (1962), Feldhusen, Ramharter, and Birt (1962), and Reed and Hayman (1962). Coulson and Silberman (1959) compared three conditions on materials taught by machine: multiple choices versus constructed responses, small steps versus large steps and branching versus no branching presentation. This ambitious program’s results indicated (1) that small steps (more items per concept) result in higher scores, but more training time, (2) the branching versions were not significantly different, but when time and amount of learning, the differences favored the branching version, and (3) there was no significant difference in the results of the type of response.

### 2.2.21 Error Rate

A major tenet in programmed instruction was presenting a sequence of instruction, which has a “high probability of eliciting desired performance” (Taber et al., p. 169). This sequence can sometimes be made too easy or too difficult. Error Rate
is associated closely with size of step because of the co dependence of the two. Skinner’s (1954) thesis is that errors have no place in an effective program. They hinder learning. Others feel it is not necessarily an easy program (with few errors) that allows more learning but the program that involves and stimulates participation. Again the results are mixed and generally dependent upon the situation. Studies by Keisler (1959), Meyer (1960), and Holland and Porter (1961) support the concept of low error rate. While Gagne and Dick (1962) found low correlations between error rate and learning others found the specific situation, topic programmed Instruction, or content to be a major factor in this determination. Goldbeck and Campbell (1962) found overt responses were less effective in easy programs. Melaragno (1960) found that when errors occurred in close proximity in the program there was a negative outcome in achievement.

Several studies have looked at the question of the use of explanations for wrong answers. Bryan and Rigney (1956) and Bryan and Schuster (1959) found that explanations were particularly valuable with complex data. However, Coulson, Estavan, Melaragno, and Silberman (1962) found no difference in achievement between a group using linear programs with no knowledge of errors and a group using branching programs that provided explanations of errors. However, the students’ level of understanding increased with explanation of errors.

2.2.22 Program Influence by Age or Level

Glaser, Reynolds, and Fullick (1963; as cited in Taber et al., 1965) conducted an extensive research study on program influence by grade level. This study was conducted within a school system using programmed materials at various grade levels, including first grade math, and fourth grade math subjects. The results were measured by program tests, teacher made tests and by national standardized tests. One purpose of this study was to determine if very young students could work on and learn from programmed materials in a day-by-day plan. Glaser et al. reported that the students were successful in learning from the programmed materials, that students who completed the programs in the shortest time did not necessarily score the highest, that 95 percent of the students achieved 75 percent subject mastery, and 65 percent of the students at the fourth grade level achieved 90 percent on the program and standardized test. While the researchers felt that the study was a success, they still felt that the role of the teacher insured proficiency by the students. Many studies were conducted in the
business and industry sector dealing with programmed instruction for training and reported significant instructional training success, a significant saving of time, or both (Hain & Holder, 1962; Hickey, 1962; Holt, 1963; Hughes & McNamara, 1961; Lysaught, 1962). A series of studies (e.g., Dodd, 1967; Evans, 1975; Mackie, 1975; Stewart & Chown, 1965) reviewed by Hartley and Davies (1978), concentrated on adults’ use of programmed instruction. They concluded that there was no single best form (e.g., format, type) of programmed instruction, which is “appropriate for everyone at a given age doing a specific task” (p. 169). They also concluded that adults like and will work with programs longer than younger students and the more interaction built in, the more it is accepted by the adults.

2.2.23 Type of Response - Constructed vs. Multiple Choices

When errors (what some call negative knowledge) are made in a program in Skinner’s (1958) view inappropriate behavior probably has occurred. Effective multiple-choice questions must contain opportunity for wrong answers and thus is out of place in the process of shaped programming Instruction behavior. Pressey (1960) and others claimed just the opposite, that “multiple-choice items are better because errors occur, permitting elimination of inappropriate behavior” (Holland, 1965, p. 86). Several studies (Burton & Goldbeck, 1962; Coulson & Silberman, 1960; Hough, 1962; Price, 1962; Roe, 1960; Williams, 1963) compared constructed response and multiple-choice responses but found no significant differences. Fry (1960) however, found constructed responses to be the better approach. Holland (1965) suggested a major advantage of programmed materials over other instructional methods that increase the probability of a correct answer. Non programmed materials generally do not require an immediate answer or response, or the material is extraneous as far as the response is concerned. The more highly programmed materials have been demonstrated to be more effective in Holland’s view.

2.2.24 Individual versus Group Uses

Several studies have been conducted to assess the value of using programmed materials (various formats) in a group setting versus individual use. The results are mixed, Keisler and McNeil (1962) reported the findings of two studies using programmed materials, one showing a significant difference favouring the individual approach over the group approach. The second study found no significant difference in group or individual approaches. Likewise, Feldhusen and Birt (1962) found no
significance between individual and group approach. On the other hand, Crist (1967), reported positive results with group work with the use of programs over individual use.

### 2.2.25 Research Concerns

As noted earlier, there has been much concern about the quality of research during the era of Programmed Instruction (Allen, 1971; Campau, 1974; Dick & Latta, 1970; Holland, 1965; Lockee et al., 2001; Lumsdaine, 1965; Moore, Wilson, & Armistead, 1986; Smith & Smith, 1966). There appears to be two major fundamental issues of concern, (1) poor research techniques and reporting, and (2) the preponderance of the comparison study. Smith and Smith (1966) noted several issues concerning Programmed Instruction research. These included: 1. Many of the comparisons used small groups, for limited subject areas and for very short study duration, 2. Because the concept of programmed instruction was relatively new in the 1950s and 1960s, the novelty effect tends to favour the new techniques, and 3. There are many uncontrolled effects apparent in many of the experiments, e.g., time.

Holland (1965) pointed out that no program or no conventional method is generic. Each program or teaching method is different in several ways (they have many, many characteristics that are uncounted for in many of these studies). The “adequacy of any method can be changed considerably by manipulating of ten subtle variables” (p. 107). Holland indicated that research on programmed learning was hampered by poor measures, test sensitivity, and experimental procedures. Campeau (1974), and Moldstad (1974) indicated rampant problems including lack of control, faulty reporting, small number of subjects, and a lack of randomization were present in many studies of this era (1950–1970). Stickell (1963) reviewed 250 comparative media studies conducted during the 1950s and 1960s and only 10 could be accurately analyzed. Most of the results were unpredictable. His general assessment has great bearing on the era of programmed instruction research. The reliance on the comparison study for much of the research published during this time illustrates examples of faulty design and interpretation. Comparison studies assumed that each medium (e.g., programmed instruction) was unique and could or could not affect learning in the same way. This medium, in the researchers’ views, was unique and had no other instructional attributes. These researchers give little thought to the medium’s characteristics or those of the learners (Allen, 1971; Lockee et al., 2001). Review of the many programmed instruction studies reveal incomplete, inaccurate, little or no descriptions of the
treatments, methodology and results (Moore, Wilson, & Armistead, 1986). Many of these studies used very small samples (if they were actually samples), lacked randomization and misused and misinterpreted results. For example, a good number of this era’s research studies used the statistical term, no significant difference to mean that variables were equally good or bad. Ask a poor question get a poor answer. Clearly any outcomes reported in these types of studies are invalid, but this fact did not stop many of the researchers, and for that matter, journal editors from misinterpreting or reporting these results (Levie & Dickie, 1973; Lockee et al., 2001).

Stolurow (1961) felt that while research indicated that learners from learning disabled students to graduate students could effectively learn from programmed devices, additional research should continue and a systematic study of programming variables be developed.

Glaser (1960) noted early on in the era of programmed learning research that “present knowledge can scarcely fail be an improvement over anachronistic methods of teaching certain subjects by lecturing to large classes” (p. 30). Even at that time there was desire to deemphasize hardware and machines. But Glaser indicated that machines had the opportunity to offer tangibility over an existing instructional method alone and programmed machines had the opportunity to showcase the capabilities of reinforcement contingencies.

In early reviews of literature, Stolurow (1961) reported three general findings on programmed learning research: (1) a programmed machine can significantly enhance learning (2) the advantages of programmed instruction are not limited by learning task or subject and (3) teaching by programs are applicable to a variety of learners.

Stolurow (1961) in his summary of programmed learning literature stated that knowledge-of-results should be studied in more detail. He felt that knowledge of results would be more effective if given earlier in a learning situation and should be a bigger factor in programmed machine and material development. While in Holland’s (1965) view, the results of programmed variables have on paper supported the general theoretical foundations of programmed learning; the research has not “improved upon the principles because the studies have been limited to gross comparisons” (p. 92). He suggested future research, including the following aspects: (1) that the measuring and specifying of variables be more exact and (2) that the research should be directed to improving existing procedures or developing programmed Instruction new techniques. The versus statements found in many comparison study titles suggest crude
dichotomies, without considering factors that might otherwise influence outcomes, such as other characteristics of the technology or the characteristics of the learner. “Consequently, a generalization of results is difficult since magnitudes of differences are important variables cannot be specified for either experimental materials or programs” (Holland, 1965, p. 92). That been said, Holland goes on to state that the research that to date (1966) supported the general principles of programming and in a paradoxal statement proclaimed that “it is perhaps comforting that comparison studies almost always show large advantages for programmed instruction” (p. 107). Holland (1965) stated that a contingent relationship between answer and content was important, that low error rate had received support, sequencing content was important and public, overt responses were important.

Hoko (1986) summarized his review of literature on the effects of automated instructional and traditional approaches, by indicating that each are unique and have specific potentials. He concluded, “The two should not be compared, but investigated, each for its own truths” (p. 18). According to Smith and Smith (1966), the most valuable aspect of the program machine and instruction literature and research as that it provided “a new objective approach to the study of meaningful learning while at the same time provides new insights into how such learning occurs” (p. 326). While much of the research on programmed learning might be described as inconclusive, contradictory or even negative, there were important contributions. These contributions included focusing attention on reinforcement learning theory and possibly its shortcomings and thus opened the possibilities of new study and experimentation. Secondly, while not necessarily the norm, there were good researchers during this time that completed solid studies that did result in significant and meaningful results. This alone should indicate a need for more variability and research control to achieve real understandings of the programming theory and methods (Smith & Smith, 1966). Some authors and researchers felt that by the middle of the 1960s changes were in order and emphasis should change from emphasizing what the learner should do to what the programmer should do (Hartley, 1974). Some educators even felt that the psychology used to justify programmed instruction was becoming restrictive (Annett, 1969). Smith and Smith (1966) and Hartley and Davies (1978) tended to believe this earlier period of programming research started to shift from looking at program variables and learner needs to dealing with interactions with entire teaching and learning systems. Smith and Smith (1966) observed that this new emphasis on “systems study will not confine its
efforts to evaluating specific machines or techniques, but will broaden its interests to include all types of classroom techniques and materials” (p. 326). Computer-assisted instruction (CAI) and computer based instruction (CBI) can be regarded as sophisticated extensions of programmed instruction theory and concept. Although some CBI research has been conducted within the context of programmed instruction, many of these studies have been conducted outside this context. Because of the many instructional possibilities that the computer can offer, many researchers consider it to be a separate field. This chapter’s literature review dealt, for the most part, only with programmed instruction regarding theory and design. It should be noted that Programmed Instruction, CBI, and CAI have similar goals to provide instruction, effectively, efficiently and hopefully economically. It is evident that the foundations of computer mediated instruction are based upon Programmed Instruction theory and research identification of measurable learning outcomes, mastery learning techniques and the evaluation of instruction. In “Programmed Instruction Revisited,” Skinner (1986) proposed that the small computer is “the ideal hardware for Programmed Instruction” (p. 110). Extending the idea of the self paced attribute of Programmed Instruction is the advent of the networked learning environment, making educational opportunities available anywhere and anytime. The revolution of the desktop computer, While trends in educational philosophy and learning theory have shifted away from behavioural sciences to more cognitive and constructivist approaches, these authors contend that Programmed Instruction has never really ceased to exist. Its influence is apparent in the instructional design processes that have continued to serve as the standards for our field (i.e., Dick, Carey, & Carey, 2000; Gagne, Briggs, & Wager, 1992; Gustafson & Branch, 1997, 2002; Kemp, Morrison, & Ross, 1998; Smith & Ragan, 1999). Recent literature regarding current trends in instructional design and technology indicates that while the systematic instructional design process has been embraced at varying levels across different venues (Reiser & Dempsey, 2002), its behavioural origins are still evident and notions of Programmed Instruction are found in existing practice. From the conduct of a needs assessment to the establishment of clearly defined and measurable objectives, to the process of task analysis, the creation of assessment instruments and approaches that reflect the specified outcomes, the provision of opportunities for practice and feedback, to evaluation of the instructional program or product all of these aspects of instructional design developed into the formation of a cohesive process as function of the Programmed Instruction movement.
Perhaps the most prominent effect of the Programmed Instruction tradition on education as a whole is the convergence of the science of learning with the practice of teaching, the point originating from the first discussion of Programmed Instruction from Skinner (1954) himself in “The Science of Learning and the Art of Teaching”.

2.2.26 Summary

In the present section, the studies under the heading ‘The Development of Self-Instructional Material’ was discussed. Under the above heading, the components in development of self-instructional material such as Teaching Machines, Psychological principles and issues, Operational Characteristics of Programmed Instruction, Linear Versus Branching Systems, Evaluation and Revision, Research on Programmed Instruction, Skinner’s Teaching Machines, Crowder’s Intrinsic Programming, Programmed Instruction Variations and Components, Ability and Individual Differences, User Attitude, Conventional Instruction (Comparison Studies), Programmed Variables (Essential Components), Mode of Presentation, Overt Versus Covert Responses, Prompting, Confirmation, Sequence, Size of Step, Error Rate, Program Influence by Age or Level, Type of Response Constructed vs. Multiple Choices. This section could help the investigator the process of development of self-instructional material and the measures could be taken for preparing a quality material.

2.3 STUDIES IN ‘DISTANCE EDUCATION - WITH REFERENCE TO INSTRUCTIONAL MATERIAL, QUALITY, CREDIBILITY’

Research on distance education has been subject to harsh and consistent critique (Berge & Mrozowski, 2001; Bernard et al., 2004; Perraton, 2000; Saba, 2000). Moore (1985) stated that there is “a massive volume of amateur, unsystematic, and badly designed research producing information of very little value” (p. 36). Panda (1992) analyzed the Indian distance education literature and concluded that “most of the studies are either descriptive status surveys or experimental studies with poor methodological footing” (p. 322). Saba (2000) criticizes the lack of theoretical underpinnings: “Research questions are rarely posed within a theoretical framework or based on its fundamental concepts and constructs” (p. 2). He is supported by Perraton (2000): “An examination of existing research shows that it is often a theoretical and predominantly descriptive” (p. 1). Is this really the case?

Lee, Driscoll, and Nelson (2004), emphasis that “understanding trends and issues in terms of topics and methods is pivotal in the advancements of research on
distance education” (p. 225). The structure of a research discipline forms the foundation for identifying gaps and priority areas (Mishra, 1998, p. 281). Based on a validated classification of research areas in distance education through the systematic analysis of expert responses in a Delphi study (Zawacki-Richter, 2009), we conducted a literature review of previous distance education research published in five prominent distance education journals between 2000 and 2008 (N = 695 articles). The aim of this paper is to address questions in the following three areas:

❖ Issues in distance education research: What are the main research areas in distance education and how are they changed between 2000 and 2008? What are the most common research areas and where are there gaps in distance education research?

❖ Research methods: Getting beyond the question of “any significant differences?” experts advocated more qualitative approaches in distance education research (Minnes, 1985; Saba, 2000). But can we observe a significant trend towards the application of more qualitative methods and mixed methods designs (triangulation)? Is there an association between research methods and gender? And do distance education journals prefer to publish qualitative or quantitative studies?

❖ Publication and authorship patterns: Researchers are more and more often involved in complex international collaborative projects. Is there a significant trend towards more collaboration among researchers in distance education? Who are the leading contributors of research papers and where do they come from? Do distance education journals tend to publish papers from their country of origin?

Based on the review of research areas and trends, the results can be used to explore the body of knowledge in distance education and to identify priority areas for future research projects.

Sample and Methods: Selection of Journals and Articles: Five journals were reviewed for this study: Open Learning (OL), Distance Education (DE), the American Journal of Distance Education (AJDE), the Journal of Distance Education (JDE), and the International Review of Research in Open and Distance Learning (IRRODL). They were selected because of their reputations as the most prominent and recognized journals in the field of distance education. With the exception of IRRODL, which was released for the first time in 2000, the journals have been used as data sources in
previous studies (cf. Berge & Mrozowski, 2001; Lee, Driscoll, & Nelson, 2004; Koble & Bunker, 1997; Mishra, 1997; Rourke & Saabo, 2002; Scriven, 1991). Furthermore, the five journals were selected because of their wide scope of distance education research in contrast to more specialized journals such as the Online Journal of Distance Learning Administration or the International Journal of Distance Education Technologies.

All articles published between 2000 and 2008 in the five journals were reviewed (N = 695). Berge & Mrozowski (2001) presented a large scale review with articles published between 1990 and 1999. Therefore, the year 2000 was chosen as the cut-off date for this study.

In order to prepare this review, an international Delphi study was carried out in 2008 to develop a classification system for research areas in distance education. The Delphi technique was selected to develop a consensus among a group of experts on common areas that are or should be covered in distance education research. The essential element in the Delphi process is anonymity of participants when giving their opinion. The Delphi method, then, alleviates problems that could be caused by domination of the group by a few prestigious or powerful individuals (Charlton, 2004). According to Isaac and Michael (1995), the Delphi method of group interaction avoids the following disadvantages of face-to-face discussions: the bandwagon tendency, the vulnerability to manipulation, and the reticence on the part of individuals to change their minds in front of others. The final expert panel comprised 25 individuals from 11 countries (Australia, Brazil, Canada, China, Fiji, Germany, Ireland, New Zealand, South Africa, UK, and USA). Based on a literature review and a qualitative analysis of the responses from the panelists, three broad meta-levels of distance education research were derived:

1. Macro level: distance education systems and theories,
2. Meso level: management, organization and technology,

Within these three levels, the research issues that are considered important by the experts can be categorized into 15 research areas. They are briefly characterized below. A detailed description of the method and the results of this pilot study is published in Zawacki-Richter (2009).
**Macro level: Distance education systems and theories.**

1. Access, equity, and ethics: The democratization of access to distance education afforded by new media and by finding ways to deliver high-quality education to those who have limited resources and poor infrastructure; issues that refer to the (sustainable) provision of distance education in developing areas. What is the impact of distance education (e.g., via mobile learning) on narrowing the digital divide and what is the role of ICT (information and communication technologies) and/or OER (open educational resources) in terms of access to education?

2. Globalization of education and cross-cultural aspects: Aspects that refer to the global external environment and drivers, the development of the global distance education market, teaching and learning in mediated global environments, and the implications for professional development.

3. Distance teaching systems and institutions: Distance education delivery systems, the role of institutional partnerships in developing transnational programmes, and the impact of ICT on the convergence of conventional education and distance education institutions (hybrid or mixed-mode).

4. Theories and models: Theoretical frameworks for and foundations of distance education, e.g., the theoretical basis of instructional models, knowledge construction, interaction between learners, or the impact of social constructivism learning theories on distance education practice.

5. Research methods in distance education and knowledge transfer: Methodological considerations, the impact of distance education research and writing on practice, and the role of professional associations in improving practice. Literature reviews and works on the history of distance education are also subsumed within this area.

**Meso level: Management, organization and technology.**

1. Management and organization: Strategies, administration, and organizational infrastructures and frameworks for the development, implementation, and sustainable delivery of distance education programmes. What is required for successful leadership in distance education? Distance education and policies relating to continuing education, lifelong learning, and the impact of online learning on institutional policies as well as legal issues (copyright and intellectual property).

2. Costs and benefits: Aspects that refer to financial management, costing, pricing, and business models in distance education. Efficiency: What is the return on
investment or impact of distance education programmes? What is the impact of ICT on the costing models and the scalability of distance education delivery? How can cost effective but meaningful learner support be provided?

3. **Educational technology:** New trends in educational technology for distance education (e.g., Web 2.0 applications or mobile learning) and the benefits and challenges of using OERs, media selection (e.g., synchronous vs. asynchronous media), technical infrastructure and equipment for online learning environments, and their opportunities for teaching and learning.

4. **Innovation and change:** Issues that refer to educational innovation with new media and measures to support and facilitate change in institutions (e.g., incentive systems for faculty, aspects referring to staff workloads, promotion and tenure).

5. **Professional development and faculty support:** Professional development and faculty support services as a prerequisite for innovation and change. What are the competencies of online teachers and how can they be developed?

6. **Learner support services:** The infrastructure for and organisation of learner support systems (from information and counselling for prospective students about library services and technical support to career services and alumni networks).

7. **Quality assurance:** Issues that refer to accreditation and quality standards in distance education. The impact of quality assurance and high quality learner support on enrolments and drop-out/retention, as well as reputation and acceptance of distance education as a valid form of educational provision.

**Micro level: Teaching and learning in distance education.**

1. **Instructional design:** Issues that refer to the stages of the instructional design process for curriculum and course development. Special emphasis is placed on pedagogical approaches for tutoring online (scaffolding), the design of (culturally appropriate) study material, opportunities provided by new developments in educational technology for teaching and learning (e.g. Web 2.0 applications and mobile devices), as well as assessment practices in distance education.

2. **Interaction and communication in learning communities:** Closely related to instructional design considerations is course design that fosters (online) articulation, interaction, reflection, and collaboration throughout the learning and teaching process. Special areas include the development of online
communities, gender differences, and cross-cultural aspects in online communication.

3. Learner characteristics: The aims and goals of adult learners, the socio-economic background of distance education students, their different learning styles, critical thinking dispositions, and special needs. How do students learn online (learner behavior patterns, learning styles) and what competencies are needed for distance learning.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Research Area</th>
<th>Level</th>
<th>F</th>
<th>%</th>
<th>Cum. %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Interaction and communication in learning communities</td>
<td>3</td>
<td>122</td>
<td>17.6</td>
<td>17.6</td>
</tr>
<tr>
<td>2</td>
<td><strong>Instructional design</strong></td>
<td>3</td>
<td>121</td>
<td>17.4</td>
<td>35.0</td>
</tr>
<tr>
<td>3</td>
<td>Learner characteristics</td>
<td>3</td>
<td>113</td>
<td>16.3</td>
<td>51.2</td>
</tr>
<tr>
<td>4</td>
<td>Distance teaching systems and institutions</td>
<td>1</td>
<td>62</td>
<td>8.9</td>
<td>60.1</td>
</tr>
<tr>
<td>5</td>
<td>Educational technology</td>
<td>2</td>
<td>48</td>
<td>6.9</td>
<td>67.1</td>
</tr>
<tr>
<td>6</td>
<td>Quality assurance</td>
<td>2</td>
<td>41</td>
<td>5.9</td>
<td>72.9</td>
</tr>
<tr>
<td>6</td>
<td>Professional development and faculty support</td>
<td>2</td>
<td>41</td>
<td>5.9</td>
<td>78.8</td>
</tr>
<tr>
<td>7</td>
<td>Access, equity and ethics</td>
<td>1</td>
<td>31</td>
<td>4.5</td>
<td>83.3</td>
</tr>
<tr>
<td>8</td>
<td>Theories and models</td>
<td>1</td>
<td>24</td>
<td>3.5</td>
<td>86.8</td>
</tr>
<tr>
<td>9</td>
<td>Learner support services</td>
<td>2</td>
<td>23</td>
<td>3.3</td>
<td>90.1</td>
</tr>
<tr>
<td>10</td>
<td>Management and organisation</td>
<td>2</td>
<td>18</td>
<td>2.6</td>
<td>92.7</td>
</tr>
<tr>
<td>11</td>
<td>Research methods in DE and knowledge transfer</td>
<td>1</td>
<td>13</td>
<td>1.9</td>
<td>94.5</td>
</tr>
<tr>
<td>11</td>
<td>Globalization of education and cross-cultural aspects</td>
<td>1</td>
<td>13</td>
<td>1.9</td>
<td>96.4</td>
</tr>
<tr>
<td>11</td>
<td>Innovation and change</td>
<td>2</td>
<td>13</td>
<td>1.9</td>
<td>98.3</td>
</tr>
<tr>
<td>12</td>
<td>Costs and benefits</td>
<td>2</td>
<td>12</td>
<td>1.7</td>
<td>100.0</td>
</tr>
</tbody>
</table>

**Total**| 695 | 100.0 |

*Level: 1=macro; 2=meso; 3=micro; F=frequency; Cum. %=cumulative %*
Table 2.2: Frequency of Articles by Research Area between 2000 and 2008 (N = 695)

<table>
<thead>
<tr>
<th>R. area*</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>7</td>
<td>11</td>
<td>31</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>13</td>
<td>15</td>
<td>1</td>
<td>3</td>
<td>9</td>
<td>3</td>
<td>8</td>
<td>9</td>
<td>62</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>24</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>8</td>
<td>4</td>
<td>10</td>
<td>7</td>
<td>48</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>14</td>
<td>5</td>
<td>2</td>
<td>41</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>12</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>41</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>11</td>
<td>15</td>
<td>12</td>
<td>10</td>
<td>8</td>
<td>22</td>
<td>14</td>
<td>15</td>
<td>121</td>
</tr>
<tr>
<td>14</td>
<td>7</td>
<td>12</td>
<td>14</td>
<td>17</td>
<td>25</td>
<td>13</td>
<td>8</td>
<td>15</td>
<td>11</td>
<td>122</td>
</tr>
<tr>
<td>15</td>
<td>17</td>
<td>12</td>
<td>13</td>
<td>11</td>
<td>13</td>
<td>8</td>
<td>19</td>
<td>6</td>
<td>14</td>
<td>113</td>
</tr>
<tr>
<td>Totals</td>
<td>72</td>
<td>69</td>
<td>87</td>
<td>62</td>
<td>76</td>
<td>81</td>
<td>82</td>
<td>80</td>
<td>86</td>
<td>695</td>
</tr>
</tbody>
</table>

*1=Access, equity and ethics; 2=Globalization of education and cross-cultural aspects; 3=Distance teaching systems and institutions; 4=Theories and models; 5=Research methods in distance education and knowledge transfer; 6=Management and organization; 7=Costs and benefits; 8=Educational technology; 9=Innovation and change; 10=Professional development and faculty support; 11=Learner support services; 12=Quality assurance; 13/Instructional design; 14=Interaction and communication in learning communities; 15=Learner characteristics

2.3.1 Quality Distance Education: The Concept

In higher education in general and distance learning in particular quality issues are becoming more important than ever before. Quality in distance learning is a multidimensional concept defined in various ways by many educationists. There are no prescriptive tools or international standards to measure quality. Yet quality is viral for the survival of distance learning programmes and is the biggest issue facing us in the years to come.

In the context of distance learning, what constitutes quality characteristics can be thought of as high quality instructional material, efficient and effective academic support, logistics of students receiving right time and a strong research base, of the many issues involved, this section focuses on the studies on distance education with reference to quality of instructional material?

The quality of the print may be reflected by validity in terms of its content, ability to understand by the student learner and readability in terms of being learner
friendly. It should be designed to be not only comprehensive but also interacting and motivating enough to compensate for the absence of classroom teaching. The learner should be able to learn on his own without a teacher and able to develop skills through self directed learning.

The preparation of such material not only requires personnel who are competent to do so but also requires much time and financial resources.

Assurance and sustenance of quality in education is a complex phenomenon. Different educationalists have perceived quality education in different ways. Judgment about quality differs according to whose views are sought (Robinson, 1992). Indian complex socio-economic system as much as access impacts quality and access, so does a quality impact access (Premji 2003). Thus, quality of education means quality of teachers, quality of learners, quality of courses, quality of planning and management, quality of infrastructure/resources and quality of teaching and evaluation methods. In terms of the system approach the quality of input and process will shape the quality of outputs that the graduates of the education system. Therefore, the quality of distance education is not a one-time affair. It is a continuous process involving sustained efforts. Open Distance Learning (ODL) in the present form is an emerging mode of imparting knowledge, skills and attitudes to learners in a non-contiguous situation. Over the past few years, there has been increasing interest in quality in ODL. The goal of ODL is to empower the learners by facilitating their learning in a high quality and learner-centered environment. Openness in education is a feature that has attracted wider support from and interest among people in many countries in recent times. Openness, a broad and relative term, is counted in terms of flexibility in the educational system: relaxed admission requirements (most obviously liberal entry qualifications), age, pace of learning, course combinations, selection of objectives, mode of instruction, evaluation techniques and so on. Education at one’s home or work place is an important feature of open distance learning (ODL). Advancements in pedagogy / instructional design and communication technologies have added to the effectiveness and efficiency of the openness in the ODL system. Information and communication technologies (ICTs) have provided various options of imparting education which is an essential component of any system to succeed. As a result, distance educators today are equipped with a variety of means and methods of education and training to make the education resources accessible to all those who want to have it as per their needs and convenience. Distance education
institutions in any country are established with the main objective to democratize education as a resource and provide every citizen, irrespective of sex, caste and creed, easy and affordable access to quality education. It is desirable for the distance education institute to proceed by maintaining high quality standard. Quality dimension of ODL system is more complex as it requires the integration of a large number of activities, processes and operations. Various academic and administrative units are involved in the teaching, learning and evaluation. ODL system, therefore, has to commit itself in maintaining high quality in pedagogy, content as well as learner support services.

Since instructional materials in the form of study guides or modules are the equivalent of lectures in conventional universities, quality control measures used in their design, development and production should be considered against parallel measures used by lecturers in the preparation of lectures and lecture notes in conventional universities. Unfortunately lecture notes are, by and large, personal, not open to public scrutiny, and have no known and readily monitored quality control measures. Measures taken in the preparation of distance learning materials ensure that high quality products are made available to learners. Learning materials are produced by either in-house or consultant writers. In either case, quality is the hallmark. Where learning materials are written in-house, as is commonly practiced at the Open University, United Kingdom (OUUK), the materials are produced by teams of writers instead of one writer. This requirement promotes quality products through collaborative work of several experts who negotiate and harmonise their expertise in a given discipline. Quality control processes are in-built in team work because team members vet each other's contribution for academic soundness, relevance, pedagogic and andragogic value and other quality indicators. The quality of such collaborative team work is usually superior to any lecture notes prepared by one content expert in conventional systems. In cases where learning materials are produced by external consultants, these are often recruited from among the best professors in conventional universities. There is financial incentive for them to produce quality materials, unlike conventional lecturers, who do not receive additional emoluments for producing lecture notes. The curriculum is produced by some of the best subject experts while the materials are written following distance learning principles and guidelines provided by specialist instructional designers. Particular attention is taken in the design and
production of learning materials because the quality of the educational process in distance education depends largely on the quality of learning materials.

The quality of distance learning materials can be judged using several indicators. They can be judged intuitively, that is, whether or not they ‘feel right’. Secondly, they can be judged informally on the basis of their adoption for use by other institutions or students within conventional institutions. A more formalized way suggested by Rumble (1986, 208) requires the system to obtain feedback from those using the materials, both students and tutors, and on the utility of the materials and the difficulties which they encountered in using them. Routine feedback can obtain a great deal of information on, for example, the amount of time that students spent on the elements of the course, whether or not they are on schedule or behind schedule on paced courses, and on what they felt about the material. Tutors may report on any difficulties which students have in understanding the material. An examination of the perceptions of distance education students, what constitutes quality in distance education materials shows that, students’ preferences are in complete agreement with the underlining principles and guidelines used by instructional designers when preparing distance study materials. According to Viljoen, Holt and Petzall (1991,12) distance education students on a Deakin University MBA programme reported that good in-text questions, accompanied as far as possible by suggested answers, materially enhances the quality of study guides. Such questions and answers enable students to test their understanding of the materials and serve as a substitute, to some extent, for the relative lack of face-to-face contact with lecturers.

According to Holmberg (1985), when distance education students interact with learning materials, they are engaged in simulated dialogue with their lecturers; hence quality distance education materials should be thought-provoking and challenging. The materials should be well written and easy to read. Well designed distance education learning materials should promote creative thinking and critical analysis rather than inhibit it. According to Viljoen, Holt and Petzall (1991, 13), Students not only value interaction with appropriately designed learning materials, but they also value the way in which the learning materials stimulate opportunities for further interaction outside the text itself (i.e. interaction with the workplace, the study group, and in residential schools, as well as ongoing interaction with teaching staff). It is important periodically to evaluate the perceptions of students and teaching staff on learning materials in order that quality materials are produced.
Research in Distance Education has been carried out on a wide range of issues. On careful examination, the studies documented so far indicate certain patterns. On the basis of the emerging patterns, researches in Distance Education have been classified into the following headings:

- Growth of Distance Education: Policy, Planning and Managements
- Programmes and Courses
- Distance Learners
- Instructional Processes
- Course Development and Evaluation
- Output and Impact of Distance Education
- Economics of Distance Education

The present study confined to instructional process and Course development and evaluation with reference of quality and credibility.

Distance education depends, to great extent, on educational technology. As means open learning, the functioning of distance education depends on better utilization of instructional strategies. In other words, distance education institutions on production of quality learning material, use of a wide range of media emphasizing varieties of learning opportunities, and properly organized face to face contact programmes. It is noted that major instructional activities of distance education have been restricted to print material and occasional interaction between tutor and students through assignments and personnel contact programmes.

It has been noticed that all the distance education institutions at university as well as the school level depend mainly on print-based instructional material. They may be in modular forms, in self instructional programmed learning packages, in semi programmed form or in some other forms as prescribed by respective institutions (Biswal, 1979; Singh, 1980; Khan, 1982; Singh 1983, Sahoo, 1985; UGC, 1986). On the part of students, it was found that a large majority of them at the university level (81 to 85 percent) depend on printed lessons for completing their studies (Anand, 1979; Khan, 1982; Sahoo, 1985). However at the school level, around 79 percent of the students depended on regular text books, notes digestes besides study material (Singh, 1983).

All the studies reported the usefulness of printed study material as perceived by the students in one or other respect (Anand, 1979; Biswal, 1979; Singh, 1980; Khan, 1982; Pillai and Mohan, 1982; Kumar et al 1986; UGC, 1986).
cases, study materials were not available in respective distance education institutions and did not follow a systematic format. While most of the students, dropouts and products of distance education expressed moderate views about the different aspects of lessons like style presentations, content clarity, suggested references and language (Sahoo, 1985 and UGC, 1986), majority of them pointed out their difficulties in studying the lesson (Singh, 1980 and 1989; Dewal, 1982; Koul, 1982; Khan, 1982; Sahoo, 1985; UGC 19876). A large number of students (41%) of six universities found the material too heavily packed with information causing difficulties in studying all the lessons. Moreover, irregular dispatch of lessons created problems for students (Singh, 1980; Nagaraju, 1982; Koul, 1982; Singh, 1983, Sahoo, 1985; UGC, 1986). It was suggested by students and teachers that study materials be made self instructional, reviewed appropriately, edited and proof read properly, dispatched in time along with references materials (Khan, 1982; Sahoo, 1985; Kumar et al. 1986).

Further, it was found that the students preferred lessons written in Indian languages and most teachers emphasized on the need for acquiring necessary experience through formal training, orientations and workshops for developing effective materials (Khan, 1982; Sahoo, 1985; and Koul, 1986). Mishra and Gaba (1999) conducted an exploratory study into the “use of activities in Self instructional materials by Distance Learner”. Study indicates that students make use of activities which are given in Self Learning Material. They also use white space given in the side margin. Majority of respondents’ preferred descriptive type of Short Answer Questions.

2.3.2 Course Development and Valuation

Vydehi (1984) evaluated the presentation of first year degree General English course of S.V.University, Thirupathi with regard to satisfaction of students need, attainment of the objectives of courses and instructional process and evaluation procedures adopted. Analysis of the different objectives for teaching English, methods of instruction, reactions of students and teachers and observations of personal Contact Programmes were means to identify weakness of the present structure of curriculum. An alternative student-active instructional format and a new type of distance teaching materials were prepared and used in an experimental situation. A comparison of the achievements of students following conventional and the modified material revealed results in favour of the modified approach. A similar kind of study was conducted by Sarwal (1984) at the Central Institute of English and Foreign Languages, Hyderabad
for preparation of teacher training correspondence course units for English Language Teaching. It was found that 15-week programmes were more effective than the 4-week programmes with regard to the degree of comprehension, degree of interest and degree of understanding achieved by students.

Evaluation of distance education is necessary from the viewpoint of justifying internal as well as external validity of the system. Since distance education is a society sponsored institution and aims at maximum involvement of people belonging to different cross sections of the society, everybody would rise questions and seek answers with regard to the efficiency and effectiveness of the system from one angle or the other. The criteria for evaluation may differ depending on the context of the study. For instance while social responsiveness of distance education can be considered as criterion for evaluation, the instructional comparability of distance education system with regular courses in terms of its products can be criterion. Some may be interested in studying the effectiveness of the distance education instructional system with reference to the achievement of different objectives meant for specific courses; in other cases, the evaluation criteria may be fixed with differential levels of long term impacts of distance education on its products, and match them with manpower requirements at a given time and cost effectiveness. If the system does not achieve success in a specific case, it may be worthwhile to raise questions about its limitations on the specific concerned.

2.3.3 Summary

In the present section, the studies under the heading ‘Distance Education - with reference to Instructional Material, Quality, Credibility’ was discussed. Research on distance education has been subject to harsh and consistent critique (Berge & Mrozowski, 2001; Bernard et al., 2004; Perraton, 2000; Saba, 2000). Moore (1985) stated that there is “a massive volume of amateur, unsystematic, and badly designed research producing information of very little value” (p. 36). Panda (1992) analyzed the Indian distance education literature and concluded that “most of the studies are either descriptive status surveys or experimental studies with poor methodological footing” (p. 322). Saba (2000) criticizes the lack of theoretical underpinnings: “Research questions are rarely posed within a theoretical framework or based on its fundamental concepts and constructs”. Delphi study (Zawacki-Richter, 2009) would explain the research studies on distance education between 2000 and 2008.
Classification of research areas in distance education through the systematic analysis was presented in the present section. Quality concern most of the studies were conducted on printed material and less number of studies were conducted on validation of the course.

2.4 CONCLUSION

Much of the research was conducted on instructional material for regular students, design of the studies are programmed text, linear and branching styles, content presented in between horizontal lines. Some studies are on modular approach, some are on teaching machines. Majority of the studies are very similar to Skinner type programmed instruction model.

After reviewing the research studies, at national and international scenarios, in last three sections of this chapter, it has been observed that less number of studies were conducted in the development of foundation course in general, and particularly no study was found in mathematics for distance learner of Open University. This was the major research gap, and hence, this problem of developing a foundation course in mathematics to fill this gap, which is discussed in the succeeding chapter. However, the review of the literature helped the researcher, in designing and developing the self-instructional material systematically.