THE REVIEW PROCESS IN FORMATIVE EVALUATION OF INSTRUCTIONAL TEXT: THE ROLE OF CONTENT EXPERTS AND INSTRUCTIONAL DESIGNERS

by

ALENOUSH SAROYAN-FARIVAR

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR IN PHILOSOPHY

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY
AND
COUNSELLING

McGILL UNIVERSITY

MONTREAL

ILENOUSH SAROYAN. FIRNAR
ID # 842680

FORMATIVE EVALUATION BY CONTENT EXPERTS AND INSTRUCTIONAL DESIGNERS

ACKNOWLEDGEMENTS

I am grateful for this opportunity, however brief, to express my deepest appreciation to all of those who have assisted me during an exhilarating, and at times frustrating period.

I extend my sincere gratitude to George Geis, who has skillfully guided me along this lengthy path, even from long distance, and to Bob Bracewell and Harold Stolovitch for their valuable comments during the development of the thesis.

I would like to thank Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (FCAR), which provided partial funding for this endeavor through a personal scholarship and a team grant to George Geis (290-10).

I am thankful to Vivian and Tony Mark for their editorial assistance, to Carolyn Kato for her help in the transcription and coding of the data, to Denis Bédard and Sophie Pelletier for the translation of the Résumé, and to Cynthia Weston who made me see the reality of deadlines.

I am greatly thankful to the Dental Hygiene Faculty at John Abbot College, and all the other 'Experts' who graciously participated in this study.

Finally, I owe my deepest appreciation to my husband, Behrooz Farivar, without whose encouragement and patience, I would have never made it.

ABSTRACT

This study explores and describes the processes of formative evaluation as carried out by content experts and instructional designers. It assumes that formative evaluation is an ill-defined, complex, problem solving task. Six experts (three Content Experts and three Instructional Designers), participated in this descriptive study. Subjects reviewed and revised a unit from a draft version of a self-instructional module on microbiology, while thinking aloud. Two coding schemes were developed and applied to the think-aloud protocols. Overall inter-coder reliability exceeded 89%. Qualitative data were used to describe the processes of formative evaluation, convergence patterns, and the degree of specificity of comments across subjects. Results suggest that there were between group differences in task representation, in the employed strategies, and in features of the text which were commented upon more frequently. Within group similarities in the outcome of formative evaluation were salient on a superficial level. Within group differences were more apparent when comments were compared qualitatively.

RESUME

La presente recherche explore et décrit les processus employés par des experts de contenu et des concepteurs de materiel pédagogique, lorsqu'ils effectuent des évaluations formatives. Elle se base sur la presupposition que l'évaluation formative peut être considérée comme une tâche de resolution de problemes complexe et definie de façon imprécise.

Six sujets-experts (trois experts de contenu et trois concepteurs de materiel pedagogique) ont participe a cette etude descriptive. Les sujets ont revu et revise une section d'une version brouillon d'un module en microbiologie, tout en pensant a voix haute. Deux systèmes de codification ont eté développes et utilises pour l'analyse des protocoles. Le pourcentage de fiabilité entre codeurs indépendants dépasse 89%. Des donnees de nature qualitative furent utilisées pour décrire les processus employes lors de l'évaluation formative, les patrons de convergence et le degre de spécificite des commentaires faits par les différents sujets. Les résultats obtenus suggerent des différences entre les groupes quant à l'interpretation de la nature de la tâche, quant à l'utilisation des stratégies d'évaluation formative et quant aux aspects du texte sur lesquels le plus de commentaires ont été fait. Pour chacun des groupes, des similarités évidentes furent revélées par l'analyse superficielle des donnees obtenues lors de l'évaluation formative. Des analyses plus détaillées de ces donnees montrèrent toutfois des différences au niveau qualitatif.

TABLE OF CONTENTS

CHAPTER III
METHODOLOGY

CHAPTER IV
RESULTS AND DISCUSSION

LIST OF TABLES

LIST OF FIGURES

FIGURE **PAGE**

CHAPTER I

Introduction

Formative evaluation is a component of the systems approach model for the design, development, implementation, and evaluation of educational products. It mandates that prototype instructional materials undergo an iterative process of testing (or review), and revision until they arrive at a level of quality deemed satisfactory for use (Abedor, 1971; Baker, 1970; Dick & Carey, 1985; Gagne & Briggs, 1979; Markle, 1967). In contrast to summative evaluation which is conducted after a product has been finalized, formative evaluation is intended to be carried out when the product is still in fluid form.

Several aspects of formative evaluation, in particular the sources which can provide feedback for this purpose, have been explored empirically. For instance, it has been found that draft instructional materials can be tested with a sample representing the potential users, or reviewed by various types of experts. Either source can produce a considerable amount of input or data useful for revision. The data, once translated into revision, will render the materials more effective (Abedor, 1971; Bamberg, 1978; Beach, 1979; Faigley, 1984; Rosen, 1968; Dupont & Stolovitch, 1983).

Empirical research in formative evaluation has been product oriented, with a general focus on the learner as the source of feedback for revision. For example, studies have explored the nature of quantitative and, to a lesser extent, qualitative data generated by various numbers of learners (Baghdadi, 1980; Burt, 1989). Studies have also compared the impact of roles (e.g., active or passive) assumed by the experimenter in soliciting feedback from the learner (Abedor, 1971; Carroll, 1988; Dick, 1968; Geis, 1988). Investigations which have included experts (e.g., instructional designers) in their design, have compared the outcome of formative evaluation by controlling for variables such as the use of a revision model or editing guidelines, and the skill and intuition of the instructional designer (Dupont & Stolovitch, 1983; Golas, 1982, Martelli, 1979).

While numerous kinds of experts can perform formative evaluation (Stolovitch, 1983), two types which are more frequently involved in this activity are content experts (who also teach the subject), and instructional designers. It has generally been assumed that for the instructional designer, formative evaluation includes diagnosing deficiencies in the design, presentation and organization of content. This may occur even though an instructional designer may lack specific knowledge about the subject matter. It has also been assumed that for the content expert, formative evaluation is limited to ascertaining

content accuracy or the appropriateness of the material for the target audience. Previous research, however, has not verified these assumptions, neither has it explored the processes of formative evaluation as carried out by content experts and instructional designers. The paucity of observational data on the processes of formative evaluation, and the absence of a theoretical framework based on which these processes could be defined is indicative that this area has not received adequate attention by researchers.

This study was conducted to address these gaps in the literature. Specifically, it attempted to identify the processes which comprise reviewing during formative evaluation, and the strategies employed by content experts and instructional designers during this process. These features were explored during an experimental task in which three experts of each type were asked to evaluate formatively a prototype unit from a self-instructional module on microbiology.

The Research Questions

The general question which this study addressed was as follows: What are the processes which content experts and instructional designer engage in during formative evaluation, and how are these processes similar and different between and within the two groups?

From the perspective of this research, the question

addressed several issues:

1- What are the processes present in formative evaluation?

2- How do the processes of formative evaluation compare quantitatively and qualitatively between and within content expert and instructional designer groups?

3- What is the degree of convergence on relevant processes of formative evaluation between and within content expert and instructional designer groups?

4- What strategies are employed by content experts and instructional designers to carry out formative evaluation?

In order to answer these questions, the think-aloud method (Ericsson and Simon, 1984) was utilized to capture the process of formative evaluation by each subject. Data were coded according to a coding scheme which was developed for this study and which drew upon the terminology of Human Problem Solving (Newell and Simon, 1972). Quantitative and qualitative comparisons of coded categories were then carried out between and within groups of experts.

Contributions to Knowledge

The benefits of this study can be summarized under four headings: theoretical, methodological, empirical, and practical.

Theoretical

Linking the performance of instructional design and content experts to a powerful model such as Problem Solving provides a novel framework for describing and predicting the performance of these two types of experts during formative evaluation. Until now, the revision behavior of instructional designers has been explained in terms of the revised product, and the individual skill, knowledge, and discretion of the expert reviewer. The use of an appropriate theoretical framework is a more reliable means by which a blueprint of a new model of the processes of formative evaluation could be built.

Methodological

The methodology generally associated with research on formative evaluation has been aimed at assessing the effectiveness of the materials by comparing pre- and post test scores of students on original and revised materials. At times, the data have been complemented with qualitative data from attitude questionnaires and debriefing interviews.

The application of the think-aloud method of data collection in the investigation of the processes of formative evaluation is an original endeavor. By placing formative evaluation in a problem solving context, it has become feasible to utilize a method which has generally been associated with the study of cognitive processes of individuals during a problem solving task. These two innovative applications have extended research on formative

evaluation from an analysis and comparison of outcome or product to an investigation of processes. These initial steps in qualitative research on formative evaluation may prove timely as the trend in research in the social sciences moves towards the integration of quantitative and qualitative methods.

Empirical

Scientific research is conducted to test models and hypotheses, or to build and advance theories and models. The present study contributes to both of these aspects. Detailed examination of content experts' and instructional designers' performance during formative evaluation provide the foundation for building a process model of formative evaluation. Data generated from the present study also allow identified processes in formative evaluation to be compared with the revision processes of expert writers' (Hayes, Flower, Schriver, Stratman, & Carey, 1987). Although the present study does not attempt to test the Process Model of Revision (Hayes et al., 1987) for fit, it does contribute to the pool of data on the review and revision processes from two novel sources. In doing this, it also integrates research from an applied area such as instructional design with relevant research in the area of cognitive psychology.

Practical

The immediate application of theoretical research is often not apparent to the practitioner. However, a study such as the current one can make a major contribution to the

decision making process of the practitioner. The detailed qualitative data of this study can shed light on many questionable assumptions which practitioners have been forced to make. For instance, as a result of this study, it is clearer as to what actually happens during formative evaluation. Furthermore, the study reveals that although on a superficial level there is consistency within content expert and instructional designer groups, on a deeper level, similarities diminish, and idiosyncratic practices become more prevalent. Minimally, these findings offer the practitioner a clearer set of criteria for the selection and inclusion of one or both types of experts in a formative evaluation project. Ideally, they emphasize the value of content and instructional design experts in formative evaluation.

Limitations of Scope and Key Assumptions

As described earlier, the main purpose of this investigation was to delineate the processes of formative evaluation as depicted by content experts and instructional designers during the revision of an instructional text. The scope of this study can further be delimited as follows:

1) The study focussed on the qualitative aspect of formative evaluation as performed by two types of experts. Due to the limited number of subjects in each group,

quantitative analysis was restricted to descriptive statistics only. Nonetheless, high inter-coder reliability, and consistent trends among group members allowed for certain generalizations.

2) The issue of whether the revision of one type of expert is superior to the other was beyond the scope of this study; no attempt was made to measure or compare the effectiveness of the revised outcome by the two different experts.

3) Similarly, no attempt was made to establish the superiority of feedback from experts over those from learners.

4) The stimulus text used for the experimental task was a self-instructional module in printed text. Findings, necessarily, are limited to this medium of instruction.

5) Finally, this research was based on two theoretical assumptions: a) that formative evaluation is an "ill-defined" problem solving task (problem solving as defined by Newell & Simon, 1972); and b) that writing is a cognitive activity (Bracewell, 1980), and so is its sub-process, revision.

CHAPTER II

REVIEW OF THE LITERATURE

Overview

The central core of this research topic is formative evaluation, one of the two kinds of evaluation in instructional development models. The importance of this concept is very clear. A significant amount of students' learning time is spent interacting with some type of instructional material. Reports suggest that in elementary and junior high school, this amounts to anywhere between 90%-95% of instructional time (Komoski & Woodward, 1985; Tulley, 1985). First draft materials do not generally assure mastery of content in a learner (Baker, 1970), and very few instructional materials are subjected to evaluation and revision involving a variety of sources such as learners, users, and experts, prior to their publication or use. Pflieger, Chomienne, Bordeleau, & Stolovitch (1978), conducted a survey to see what percentage of instructional materials used in K through junior high classes underwent formative evaluation in the Province of Quebec, Canada. Their findings approximate those of Komoski (1971) who conducted a similar survey in the United States: a mere one percent of instructional materials in use had undergone formative evaluation. In many of the smaller publishing

houses, it is the publishing executives who make decisions about the production, revision and marketing of instructional materials and many of their suggestions are based on economic rather than educational factors (Bowler, 1978).

The review of literature on formative evaluation provides a basic background for the study of revision carried out by instructional designers and content experts, and will therefore be presented first. As will be discussed in this review, instructional development models are heuristic systems which do not yield perfectly predictable outcomes, an attribute commonly associated with a theory. Rather than being a blueprint for the prediction of outcomes, instructional development models are procedural guidelines which are of direct and immediate assistance in a given situation. Consequently, their use in scientific research needs to be supplemented by a more rigorous theoretical framework. The general problem solving model (Newell & Simon, 1972) and the Cognitive Process Model of Revision (Hayes et al., 1987) seem promising in providing such a framework. These two models will be discussed in the latter part of this chapter with particular emphasis on the processes present in expert revisers' behavior.

The Systems Approach to the Development of Instructional Materials

Historically, instructional materials have been developed

based on various frameworks (Saettler, 1968). Friesen (1973), for instance, suggested the intuition and expertise of the master teacher and developer as one source of inspiration. Briggs (1977) made reference to the author's theoretical perspective as well as the tradition of the domain as guiding factors in materials development. Logan (1982) included 'on the spot' development in his list which implied preparing the environment and supporting the learner in whichever direction he/she chose to go. All three, however, also suggested instructional design as a systematic and reliable vehicle for developing materials.

Instructional design models (Briggs, 1977), also known as educational technology models (Henderson & Nathenson, 1977; Lumsdaine, 1966) have generally been successful in terms of defining an educational goal and designing steps to accomplish that goal. The efficiency of these models has been clearly demonstrated, for example, by their extensive application in developing training materials for the United States Army (Branson, 1973).

Recently, a number of attempts have been made to associate instructional development with a general theory of instructional design (Reigeluth, 1983; Reigeluth & Merrill, 1979; Richey, 1986). These models, however, rather than being theoretical constructs enjoy a more solid footing as heuristic systems which adhere to the principles of the systems approach in the design, implementation, and evaluation of educational materials.

A systems approach implies a well-integrated plan of operation where the output of one component is the input for the next. The approach necessitates that all parts of the system be interrelated, and aim at one common goal (Briggs, 1977; Dick & Carey, 1985). The strength of instructional development models lie in a built-in, ongoing evaluation which measures the success of each step along the way. This process is iterative, that is, it can occur at several points during the development of instructional materials, from the time of conception until the time materials appear in final form.

The nature of systematic design and development might imply that the revision activity is a redundant component since it is the function of systematic design to produce a perfect product. Nonetheless, it has been suggested that systematic design is not always a completely successful endeavor. Dick (1977) has associated the sometimes imperfect outcome of these models with inadequate underlying theories. "If the theories are weak, the product will be less than perfectly effective (p. 312)". Gropper (1975), on the other hand, has related it to imperfect execution.

To evaluate the robustness of this framework, Cowan (1980) applied the systems approach to the design of a course which did not have a syllabus or outline. He found that the model lends itself well to the refinement stage, but poorly to the development stage of course production. As Dodd, Lehunte, & Sheppard (1975) have pointed out, there are at least fifty decisions incorporated in an instructional development task,

each of which could jeopardize an otherwise perfect production. Of particular significance are those factors which are subjective in nature because they are governed by budget, time, and the specific needs of the target audience. These factors, as Gropper (1975) has also asserted, prevent the developer from a strict adherence to the model. As a result, the outcome is generally a less than perfect product. Needless to say that given these shortcomings, formative evaluation necessarily becomes a viable means by which an educational product can be further improved.

Formative Evaluation

Although there are numerous instructional development models, the components of the majority are similar to those which Dick and Carey (1985) and Gagne & Briggs (1979) have described in their widely used publications. These steps include, identifying instructional goals, conducting needs assessment, identifying entry behavior, developing performance objectives and criterion referenced tests, developing and/or selecting instructional strategies, developing and/or selecting media, conducting formative evaluation, revising instruction, and performing summative evaluation (Dick & Carey, 1985, pp. 2 & 3).

Two extensive reviews (Andrews & Goodson, 1980; Stolovitch, 1982) which have compared the components of a total

of fifty-two commonly used instructional development models demonstrate that revision, as a means of improving the product before publication, is strongly recommended (more than 98% of the time) across models. Some of the phrases used to define this recommendation include, "adjustment, revision and further evaluation" (Gagne & Briggs, 1974, p. 213); "... revision of courseware based on diagnosis" (Andrews & Goodson, 1980, p. 5); and 'product verification' (Merrill & Boutwell, 1973, p. 95).

The practice of evaluating and revising materials while they are still in draft form also seems to enjoy deep rooted precedence in the history of materials development. According to Cambre (1981), as early as 1921 a film developed for the military was evaluated so that it could be adapted for public viewing. In the late thirties and early forties, the definition of evaluation began to lend itself directly to revision for the purpose of improving materials. "Evaluation and revision" began to assume the meaning of the process of judging the value of a product and correcting its flaws at various stages of development (Cambre, 1981, p. 4).

Although a strongly recommended activity, evaluation data have not always been used to improve instructional materials. Gooler (1980) surveyed 225 evaluation studies of instructional technology which were included in Educational Resources Information Center (ERIC) between 1970 and 1978, inclusively. Only 21 of these studies stated that the primary purpose of evaluation was to compile data for revision and thus, ultimately, improve the instructional materials. Although the

author did not specify the purpose of the remaining 90% of the surveyed studies, one can assume that they were conducted in order to accept or reject a product or program.

Instructional materials have generally been subjected to evaluation for two reasons: to improve them and to decide whether or not to use them. In 1941-42, Tyler, followed by Cronbach (1963), and Scriven (1963) began to make this distinction by separating evaluation conducted during the development stage of a product and that following its completion. It is Scriven's (1967) two terms "formative" and "summative" evaluation, which have been widely used in the literature to define this distinction.

Of the two types of evaluation, summative evaluation is generally implemented after the development process is completed and is normally undertaken by an agent other than the original author/developer. It is most useful when materials are being considered for adoption or when their effectiveness in fulfilling their purpose, namely to instruct, is taken into consideration. For instance many of the incorporated revisions in new editions of textbooks have been prompted by summative evaluation.

Formative evaluation, on the other hand, is conducted expressly as a procedure "... to identify aspects of the course where revision is desirable... not in the homestretch when the developer is naturally reluctant to tear open a supposedly finished body of materials...[but] while it is still fluid" (Cronbach, 1963, p. 675).

From Scriven's (1967) publication, several characteristics of formative evaluation can be discerned. First, it is conducted after an initial first draft but before the materials are finalized and published. Second, its purpose is to improve the effectiveness of instructional content. Third, it is based on feedback received from potential users as well as experts. Fourth, it may or may not be implemented by the instructional writer / developer.

Alkin and Fitz-Gibbon (1975) have defined formative evaluation as comprising two distinct activities. One activity, labelled implementation evaluation, ascertains whether or not the product has been developed in accordance with the pre-stated purpose and intent. Another activity, progress evaluation, measures the adequacy of the prepared program by the performance and achievement of the user on criterion referenced test items. These two activities suggest that in order to carry out a formative evaluation thoroughly, it is necessary that materials be reviewed by experts as well as learners (Geis, 1987). Experts can comment on the quality and validity of the materials. Learners, by their performance, can testify whether the assumptions made about their abilities are accurate. Nonetheless, in practice, instances where both sources are used in formative evaluation are quite rare. In the next section, some reasons for the inclusion of experts in the process of formative evaluation will be presented.

Sources of Revision Data

The instructional design literature clearly recommends that materials ought to undergo evaluation both by experts (instructional designers as well as others) and potential learners. Nevertheless, in practice, formative evaluation has turned out to mean the empirical validation of materials with learners alone. This is evident in the alternating use of the term formative evaluation indiscriminately (Markle, 1976; Stakenas & Mayer, 1983) with Learner Verification and Revision (Komoski, 1983; Komoski & Woodward, 1985), and developmental testing (Geis, Burt, & Weston, 1984; Nathenson & Henderson, 1977; 1980). Both of these latter forms of evaluation exclusively use learners as the source of revision data.

The empirical research on formative evaluation has produced abundant information regarding the student as a source of data. For instance, some studies have attempted to substantiate the advantages of a revised version over an original version by using learners as the source of revision data (Abedor, 1971; Robeck, 1965). Others have compared the amount of feedback received from various number of students (e.g., one, small group, field testing) (Burt, 1989). Relying on student feedback alone is a practice which is neither recommended nor effective. As Geis (1988) has cautioned, when using students as a source for feedback, often they are given the dual role of test taker (for which they have had practice), and commentator (for which they are generally unprepared).

Yet other studies have used both learners and experts as sources of data for revision and have compared student performance on original and revised versions of the material. For instance, Rosen (1968) compared two sets of revised text, one based on performance data, that is pre- and post test scores of students, and the other based on the revision skill and intuition of the instructional designers. The original materials and the two sets of revised materials were given to students randomly assigned to three groups. The results of performance on a subsequent criterion-referenced test indicated that both revised versions were superior to the original and that revision based on empirical data was more effective than the one revised based on intuition.

While such findings may suggest that a revised version is superior to an original version, drawing generalizable conclusions regarding the superiority of the learner over the expert as a source of data, warrants careful consideration. As MacDonald-Ross (1978) has pointed out, "the equation of quality [of instructional materials] with attainment of objectives is unsound ... because it is grossly instrumental" (p. 231). Furthermore, achievement of the objectives cannot sufficiently vouch for the validity and pedagogical value of the material. As well, in the absence of a detailed account of how identified problems are translated into actual revisions, reliability between revisers remains a major issue. Studies which have investigated the effectiveness of various data gathering techniques, and sources of feedback, suffer from a similar

contaminating factor: inadequate control of the use of data in revision. As Kandaswamy (1980) has cautioned, it is highly possible that radically different modifications can be made by different revisers based on the same student performance data.

The notion of neutralizing the variability which individual revisers introduce to formative evaluation may have prompted experts, in particular instructional designers to publish revision guidelines, which in effect represent their own behavior. (See Saroyan and Geis, 1988 for a review of 48 such guidelines.) Some of these guidelines have been used in comparative studies which have substituted for the expert the use of expert generated revision guidelines. For instance, Golas (1982) investigated the cost effectiveness as well as the instructional advantages of text revised by two groups of randomly assigned instructional designers: one which revised the materials on the basis of student data, and the other based on editing guidelines. Her results indicated that in terms of learner outcome, both revisions produced identical results. However, the set which was revised using guidelines, proved to be significantly more cost effective.

There are investigators who have argued against the unequivocal preference of empirically-based tryouts and revisions over expert review or revision based on empirically proven attributes of effective instruction. MacDonald-Ross (1978), for instance, posits that revision data produced by students is seldom richer or more meaningful than changes suggested by expert revisers. Merrill, Reigeluth, & Faust

(1979) developed an instrument for the evaluation of instructional text in order to eliminate the costly process of empirical validation. However, they recommended that after revision, materials be validated with the students in order to ascertain the strength of the instrument.

One might acknowledge the fact that the expert is a cost and time effective source of feedback for revision, particularly when the alternative is using salaried trainees (Foshay, 1984). Besides the practical aspect, experts have an advantage over students in detecting erroneous or potentially problematic content, and are more likely to be capable of in the remediation of identified problems.

This however, does not suggest that the potential of students for flagging down problem areas should be underestimated. Only the learner can verify or negate the assumptions that the author/developer has made about the intended audience (Geis, 1986). Whereas expert comments are generally followed by recommendations intended to improve content, pedagogical value and presentation, the learners can successfully convey their feeling of discontent towards the material.

While there may be a lack of empirical or theoretical research in the area of formative evaluation to support assumptions about the role of experts and the processes undertaken in reviewing materials for formative evaluation, the current cognitive science literature has much to offer on this topic. Of particular relevance are investigations which have

been carried out on the topic of written composition. This body of literature has much potential to contribute towards the advancement of knowledge on formative evaluation, and the ultimate goal of rendering instructional text more effective. These possibilities will be discussed in the next section.

A Review of Related Research

Research on Writing and Revision

The relevance of research on written composition to formative evalaution is that writing is the only context within which the cognitive processes of revision have been studied in any detail. These studies represent both revision of one's own production, as well as those of others. Nonetheless, aspects of this body of literature which reveal some of the characteristics of expert revisers are relevant to formative evaluation, and will be reviewed in this section.

Converging data from various areas of research on writing suggest that composing consists of a task environment, the writer's long term memory, and the writing process (Flower & Hayes, 1981a). The latter includes planning, translating and revising. The first two components entail goal setting and the selection of production procedures, and prose generation. Revision involves the evaluation of the generated text and all subsequent activities which lead to change in that prose.

More recently, investigations have been directed toward

the detailed analysis of each of the sub-processes. For instance, Breuleux (1987a), Flower & Hayes (1981b); Hayes (1988), and Matsuhashi (1981) have looked at planning and goal setting procedures. Others (Beach, 1979; Bartlett & Scribner, 1982; Faigley & Witte, 1981; Flower, Carey, & Hayes, 1985; Flower, Hayes, Carey, Schriver, & Stratman, 1986; NAEP, 1977; Nold, 1980; Sommers, 1980; Smith, 1982; Perl, 1979) have studied revision exclusively.

One major outcome of these studies has been the rejection of the linear model of writing which was in vogue in the sixties (Rohman, 1965; Rohman & Wlecke, 1964). The current view holds that the writing sub-processes do not occur in sequence but are interactive as well as iterative. Another outcome has been the delineation of a dual function for revision. Murray (1978) has referred to it as 'internal' and 'external' revision. The former includes all of the activities which aid in the discovery of thought and are undertaken by the author to reduce discrepancy between written discourse and intent. The latter refers to all changes made in order to adhere to standard conventions of grammar and form or to improve the suitability of written text for the intended audience.

A relevant outcome of research in this area to the present study is a model of revision the processes, conceptualized on the basis of observing expert writers while they perform a revision task (Flower et al., 1985; Hayes et al., 1987). This model is based on the assumption that

writing is a problem solving activity (Hayes & Flower, 1980; Flower & Hayes, 1977), as are all of its components, including revision. Thus, regardless of whether revision is considered a phase of formative evaluation (Dick, 1980; Henderson & Nathenson, 1976b), thought discovery (Lowenthal, 1980; Murray, 1978), or editing (NAEP, 1977; Perl, 1979), it is based on the information processing theory of problem solving as defined by Newell & Simon (1972).

Information processing theories describe behavior as an interaction between an information processing system, the problem solver, and a task environment, the latter representing the task as described by the experimenter. In approaching the task, the problem solver represents the situation in terms of a problem space which is his/her way of viewing the task environment. These three components -information processing system (IPS), task environment, and problem space-establish the framework for the problem solving behavior (For a more detailed description, the reader is referred to Newell & Simon, 1972, Chapter 14).

The application of problem solving models to ill-defined tasks where the criteria for successful solution is not very clear, such as writing, has been challenged in the literature. However, there is evidence that even ill-structured problems such as composing a fugue (Reitman, 1965), or building a house (Simon, 1973), are composed of many sub-processes, each of which generally contains a well-defined goal. Conversely, studies which have investigated well-structured problems such

as theorem solving or chess (Chase & Simon, 1973; de Groot, 1966; Newell & Simon, 1972), report that the intermediate goals of well-structured problems are generally ambiguous.

As Ericsson & Simon (1984) have stated, in order to follow a subject's behavior in a problem solving task, it is essential to describe it in terms of a processing model. Such a model has been conceptualized by Hayes et al. (1987) and is based on expert writers' performance. This model may facilitate the interpretation of the formative evaluation and revision processes of instructional design and content experts. Figure 1 depicts this model.

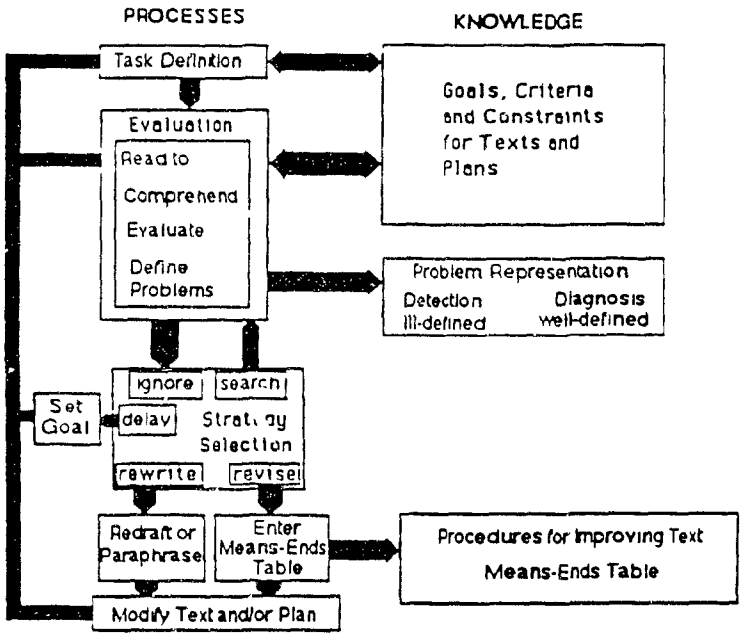

FIGURE 1. Process Model of Revision
From "Cognitive Processes in Revision" by J. R. Hayes, L. Flower, K. A. Schriver, J. F. Stratman, and L. Carey, 1987, Advances in Applied Psycholinguistics: Vol. 2, p. 185. Copyright by Cambridge University Press.

The model is composed of a set of sub-processes (presented in the left column), and knowledge states (presented in the right column) which either initiate the processes or are an outcome of them.

The first sub-process, task definition, describes a) the way in which the reviser perceives his task, b) the goals that he sets or perceives as being set for him, c) the strategies that are in his repertoire and which he will employ to achieve his goal, and d) the areas of the text on which he will focus. Task definition is dynamic and may be modified during the revision task. It is also variant across individuals, depending on set or imposed goals, criteria, and constraints.

On the same level, there are the three knowledge states of goals, criteria, and constraints. These may be self-imposed by the reviser at the outset, or may be generated and modified during the course of revision. Simon (1978) has described the same notion as a problem space. "The structure of problem space constrains behavior in a variety of ways. First it defines the legal moves. Second, it defines the goal and usually, though implicitly, the direction of the movement toward or away from the goal. Third, it interacts with the limits on short term memory to make some solution paths easier to find than others" (p. 275).

The second process, evaluation, is closely regulated by the underlying purpose of the evaluation. In other words,

depending on whether one reads the text for comprehension, evaluation or to define a problem, one may set different goals, criteria, or constraints on the task. The most significant outcome of evaluation is a problem representation which may range anywhere from a vague or ill-defined problem identification to a well-defined diagnosis of the problem. In the latter case, a recommendation for revision may also be present.

The third sub-process is strategy selection. Based on the problem representation, the reviser has one of two options: to pursue strategies which relate to the text or those which relate to the task. In the former instance, the reviser may choose to rewrite the entire text or to revise only those portions which appear to be flawed. In the latter case, the reviser may choose to ignore the task, to delay it pending the fulfillment of other conditions, or to search for additional information in order to build a clearer task definition.

The sub-processes of revision, as identified by Hayes et al. (1987) are very much in accordance with the general problem solving behavior of experts in other domains. Studies in well-defined areas such as games (e.g., de Groot, 1966), and physics (Larkin, 1983; Simon & Simon, 1978) as well as those in ill-defined areas (music composition, architecture), suggest that experts differ from novices in three general areas. These are: a) problem representation; b) pattern recognition and organization of information; and

c) selection and use of solution strategies (Anderson, 1982; de Groot, 1966; Glaser, 1985; Greeno & Simon, 1984; Larkin, 1983).

In the following section, some of the empirical research which has lent support to these theoretical assumptions about experts will be reviewed.

Research on Expert Characteristics

Glaser (1985) has posited that "relations between the structure of a knowledge base and the problem solving process are mediated through the quality of representation of the problem". More importantly, "this problem representation is constructed by the solver on the basis of domain-related knowledge and the organization of knowledge. The nature of this organization determines the quality, completeness, and coherence of the initial representation, which in turn determines the efficiency of further thinking". (p. 4).

Over the course of decades, there have been numerous attempts to classify knowledge in a given domain. For instance, Polanyi (1962) used the terms 'tacit' and 'explicit', while Anderson (1976) used 'declarative' and 'procedural' knowledge to make a distinction between knowledge of facts and the application of that knowledge in performance situations. Chiesi, Spilich, & Voss (1979)

defined knowledge of domain as "an understanding of the base concepts, [definition and relation and usage] as well as its goals, rules, and/or principles" (p. 257). Glaser (1985) proposed that experts not only possess knowledge but demonstrate rapid access to and efficient use of this body of knowledge.

Some of these theoretical postulates have been supported by empirical research. For instance, Hull's (1983) investigation of the editing processes of experts and novices indicates that experts know more rules and conventions and are better at problem solving and experimenting with text.

Other studies report that non-expert revisers, that is, students, direct their comments or revisions towards the lexical and syntactic levels of text. Experts, on the other hand, attend to semantic shortcomings, thus making changes which are more global and substantive in nature (Faigley & Witte, 1981; Nold, 1980; Perl, 1979).

In an assessment of the revision skills of nine, thirteen, and seventeen year old students, The National Assessment of Educational Progress (1977), reported that student revisers made stylistic, informational, mechanical, cosmetic, grammatical, continuational, transitional, organizational, and holistic changes with the majority of the changes clustering around the first three categories. With age, however, there was a pronounced increase in the changes made in style and the amount of informational

content.

Bridwell's (1980) results also indicate that the revisions of 12th graders are mostly mechanical corrections which are made in order to conform to the conventions of the English language. Corrections included spelling, punctuation and word selection comments.

The differences between student and expert groups may be partly due to the fact that in teaching writing to students, a great emphasis is placed on grammar and usage and students are expected to closely adhere to a socially acceptable style (Flower & Hayes, 1977). "Within the classroom 'writing' appears to be a set of rules and models for the correct arrangement of pre-existent ideas. In contrast ... in professions..., writing is a highly goal oriented, intellectual performance. It is both a strategic action and a thinking problem" (p. 229).

Metacognitive knowledge (Bracewell, 1983; Flavell, 1963) or 'intention' (Flower et al., 1986), which refers to the actual use of knowledge that one is assumed to possess, is also a prominent attribute of experts. This attribute influences both the formation of a problem representation and strategy selection. "Intention enters the process in two places: in the form of an initial problem representation - the reviser's image of the task itself- and in the form of the goals and criteria she brings to bear during evaluation" (p. 20).

The current literature suggests that experts and

non-experts define the revision task differently, each group pursuing different strategies to detect problems and make subsequent revisions. For instance, Sommers (1980) conducted a series of studies in which she compared the revision processes of experienced writers and student writers. One of her conclusions was that the students interpreted their revision task as "scratching out and doing over again", "reviewing", "marking out", and "slashing and throwing out". They considered revision as a rewording exercise, a means of curing lexical redundancy and "cleaning up speech" (p. 381). Similarly, Beacn (1979), concluded that his group of inexperienced revisers defined their task as "polishing up', more in line with copy-editing corrections rather than introducing substantive changes in content and organization.

Sommers (1980) also found that the experienced writers used words such as rewriting and revising to describe the process of revision where the primary objective was to "find form or shape of their argument (p. 384). "The writers ask: what does my essay as a whole need for form, balance, rhythm and communication... This sense, however, is constantly in flux as ideas are developed and modified; it is constantly "re-viewed" in relation to the parts (p. 386).

In studying children's revision strategies, Calkins (1980) found that they not only had difficulty in maintaining a plan of revision but had a haphazard process of detecting problems. Beach and Eaton (1984) reported that even college students had difficulty in articulating their

goals and intentions.

Similarly, Stallard (1974) compared the writing behaviour of good student writers with another group of randomly selected students. Some of the characteristics of the good writers which were not observed in the random group included an awareness of the task and an ongoing self-evaluation while performing the task. Both groups, however, demonstrated concern for spelling and mechanical flaws. Stallard (1974) also found that even his extensive student writers, like the randomly selected group, failed to show a concern about the structure of paragraphs or the essay as a 'whole'.

In making revision recommendations, another characteristic which appears to be prominent in the expert reviser is an awareness of audience needs. The underlying assumption of audience awareness lies in Piaget's concept of egocentrism which suggests that the "cognizer sees the world from a single point of view- only his own-but without the knowledge of the existence of the other viewpoints or perspectives and without the awareness that he is the prisoner of his own (Flavell, 1963, p. 60). After a series of experiments, Piaget (1926) concluded that six-year old children were very poor in adapting a message to the needs of someone other than themselves.

Cooper & Flavell (1975) have pointed out that only around grade six do children begin to develop an awareness of audience needs. However, even then, this skill has not

been ingrained deeply enough to be part of the unskilled revisers schema to be constantly present during revision. Flavell (1977), and Markman (1977) have posited that the sub-processes of planning and reviewing are late developing abilities.

A lack of concern for audience needs is not restricted to the young writer alone but appears to be a characteristic of the unskilled writer/reviser as well. Studies which have used more mature students as subjects (Crowhurst, 1978; Stallard, 1974) indicate that even college age students do not evaluate text with a continuous consideration for their reader's needs.

While the studies cited in this section generally focus on the improvement of writing and not instructional text per se, they are relevant to formative evaluation in at least two aspects:

a) The characteristics and attributes of expert revisers described above suggest that expert review is, as suggested in the literature, a significant source for formative evaluation. Learner comments is equally important although one must be aware of the fact that in most instances learners' comments are subjective preferences (Bell & Sullivan, 1981; Duchastel & Whitehead, 1980; Frase, 1981; Hartley & Trueman, 1981; Hartley, Trueman, & Burnhill, 1980).

b) They provide a theoretical framework with which the performance of content experts and instructional designers

can be compared.

Revision Techniques in Formative Evaluation

Providing feedback on the instructional text is only one aspect of formative evaluation. The other is the translation of the identified problems into actual revisions. It is usually the expert who undertakes this activity. In this section, some of the suggested techniques for revision will be reviewed.

One of the comprehensive procedural model for translating raw data into revision has been developed by Gropper (1975). Gropper's system is considered to be both a diagnostic tool which pinpoints learning and program design failures, as well as a remediation instrument which provides revision suggestions for faulty design.

This system focuses on three separate levels for incorporating revisions. These levels include individual tasks, transitions between tasks, and cumulative learning experiences. For example, if an individual task fails, the range of options available to the reviser includes the modification of the behaviour control techniques, alteration of content, and/or adjustment of language to facilitate comprehension. Likewise, if a transition proves to be too difficult to achieve or not difficult enough, the reviser is given several options: a) make adjustments in both the

quality and quantity of the content; b) alter the sequence of task types; and c) prolong exposure to the stimulus materials. Finally, if a program fails to promote cumulative learning experience, the reviser may review the needs of the target population, rearrange the sequencing of the sections, and/or check for omissions and commissions.

Despite its comprehensiveness, as Dupont and Stolovitch (1983) have pointed out, Gropper's system has not been tested for reliability. Further empirical studies need to be undertaken before it can be claimed as a successful revision tool. Dupont & Stolovitch (1983) have proceeded to adapt Gropper's discussed system, and have empirically tested it in a subsequent study. This adaptation, like its predecessor, necessitates the presence of learner data in order to be functional. For instance, to judge the adequacy of content, questions such as the degree of attainment of pre-stated objectives, the amount of congruence between content and learner prior knowledge, scope of learner interaction with materials, and the amount of transfer and recall of knowledge are posed. In their experiment, Dupont & Stolovitch (1983) used their model to revise a set of instructions on the use of a single lens reflex camera. Two groups of instructional designers revised the content: one, using the model and learner data, and the other relying on personal experience and intuition. Results demonstrated that the use of learner test scores and the model yielded more reliable revision comments than revision based on

intuition.

The limitation of both of these models remains however, in the fact that in the absence of a learner sample, one cannot use them to evaluate and revise the text on its own.

Several revision techniques have been recommended in the literature which can be carried out by experts independent of the presence of any students. These include the use of readability formulae (Klare, 1963; 1976), expert generated guidelines (Duchastel, 1983a; 1983b; Hartley, 1981; Hartley & Trueman, 1981; Merrill & Bunderson, 1981; Wright, 1977), linguistic and qualitative content analyses, and subject matter verification (Mac-Donald Ross, 1978). One must, however, consider these options in light of some of the criticisms they have received. For instance, in the case of readability formulae, it has been contended by some that despite improving readability, the reader's comprehension, and retention may not alter (Klare, 1963, p. 14). Other criticisms addressed to the use of these formulae are that they are not suitable for units larger than the sentence, and are inadequate in providing guidelines to writers on important features of writing, such as organization, emphasis or paragraph construction (Flower et al., 1980; Kniffin, 1978; Redish, 1980). In summary, although readability formulae are easy to use, and offer an objective, quantitative measure based on which revisions can be made, they fail to be effective in making structural revisions which according to some researchers (Kern, Sticht,

Welty & Hawke, 1976; Wright, 1978) are the most significant characteristic of written information.

Aside from readability formulae, content analysis, and content verification, there are other techniques which have been developed for evaluating materials by instructional designers/material developers without using student data. One such model has been promoted by the United States Army primarily for analyzing the quality of instruction but it has also been used as a tool for identifying flaws in an instructional product. This model has appeared under a variety of labels including the Instructional Quality Profile (IQP), (Merrill, Reigeluth, & Faust, 1979), the Instructional Strategy Diagnostic Profile (ISDP) (Merrill, Richards, Schmidt, & Wood, 1977), and the Instructional Quality Inventory (IQI) (Ellis & Wulfeck, 1978). Its function is based on the consistency among, and adequacy within its four linear components, namely, purpose, objectives, test, and the presentation of instruction. In other words, the quality of instruction is determined by checking a) the consistency between the purpose and objectives, b) the adequacy of the objectives, c) the consistency between the objectives and tests, d) the adequacy of the tests, e) the consistency between the tests and presentation, and f) the adequacy of the presentation. (For a more detailed overview of the Profile, the reader is referred to Choi, Merrill, Callahan, Hawkins, & Norton, 1979).

There are at least two studies which have attempted to empirically validate the ISDP. Burkholder (1981-82) used it to revise a set of self instructional materials which taught abstract concepts. The study investigated the differences that five dependent measures: learner performance, confidence, attitude, study and test time, made in three groups of comparable learners. One group of learners used an ISDP revised version of a text improved only for consistency; a second group used an ISDP revised version of text improved for both consistency and adequacy; and a third group used the original version of the materials. The results indicated a significant difference between the original and revised versions, and in particular the version revised for both consistency and adequacy on all of the dependent measures with the exception of study time. Perhaps one limitation of this study is that the validity of the ISDP has been assumed by the investigator. Given the fact that any type of revision will make the text more effective than the original version, (Abedor, 1971; Baker, 1970; Kandaswamy, Stolovitch, & Thiagarajan, 1976), one does not know how an intuitively revised version would compare with the ISDP revised version.

In another evaluative study, Choi et al. (1979) used the ISDP to rate two different organic chemistry textbooks to determine which was more appropriate for a course from an "instructional science" perspective. Their findings suggest that the textbook which received higher ratings on the ISDP,

also produced higher post-test scores on national exams. While these investigators recommend the ISDP as a more appropriate instrument for summative rather than formative evaluation, particularly for textbook selection, they also point out that it is more functional within contexts where there is a one-to-one correspondence between the text itself and test items. The emphasis of the application of these instruments is clearly on the identification of weaknesses in instructional text rather than on an explicit account of the process of translating deficiencies into revision. This aspect of formative evaluation, that is the heuristics based on which evaluation takes place, has not received much attention in the literature.

The following section will contain a review of what is known about the revision behavior of these content and instructional design experts.

Formative Evaluation by Content Experts and Instructional Designers

Many types of experts may partake in the process of instructional text development. This team work of "text assembling" (Ally, 1985; Nevo, 1977; O'Donnell, 1985; Wright, 1985) includes the following steps: the preparation of a first draft by the author and/or instructional designer; revision of this draft by a host of experts (17 in one account, Bowler, 1978) such as subject matter specialists, pedagogues, document designers, and government

and public interest groups; and the incorporation of the revision(s) into the text by the original developer or an external body such as an editor. Most often, this latter activity is based on the skill, knowledge and discretion of the reviser. Flower et al. (1980) investigated the strategies of expert writer revisers and reader revisers in the revision of functional documents. One of their conclusions was that in the absence of a set of principles to direct the revision process, documents are bound to be exposed to numerous revision styles, "each ... mak[ing] ... revision on his or her own independent criteria for 'improving' the writing".

The available literature on the prescribed or self-perceived roles of instructional design and content experts during formative evaluation, is minimal. For instance, from the instructional design literature, one can only infer that instructional designers have assumed roles ranging from a generalist implementor of the instructional design model to a coordinator and reviser of content (Faust, 1980; Nichols, 1981; Roberts, 1979). In this capacity, they have commented on design and presentation as well as on content and language issues. A survey of revision guidelines generated by instructional designers demonstrated this trend (Saroyan & Geis, 1988). When the compiled items (1009 items from 48 lists) were sorted, 445 items pertained to instructional design, 373 pertained to presentation, and 191 to content.

Details about the revision process of these experts during formative evaluation is non-existent. A decade ago, Wright (1978) suggested that by analyzing qualitative data of expert solutions, we may begin to understand the as yet undefined principles that underlie successful performance. The considerable amount of current research on the processes of revision carried out by writing experts, i.e., writing teachers (pedagogues), and professional writers and editors (content experts) can serve as an excellent model for studying the process of formative evaluation by instructional designers and content experts.

Conclusion

From the review of the literature, it can be concluded that the expert contingent in formative evaluation has not received as much attention from researchers in the area of instructional design as have learners. This paucity of research has created a vague notion about the processes of formative evaluation and the way in which identified problems are converted into revisions. While assumptions can be made regarding the outcome of formative evaluation by content experts and instructional designers, it is not evident whether the same processes and strategies are utilized by these two expert groups. On the other hand, the cognitive psychology literature, particularly studies on

written composition and expert-novice differences in revision, has described certain expert behavior patterns. These include differences in problem representation and solution strategy selection. With this information, a qualitative study of instructional designer and content expert's behavior in formative evaluation is timely. This line of research appears to be especially promising with the novel application of a general problem solving model as a theoretical framework and the use of a methodology which has proven to be effective in studying ill-defined, complex tasks.

CHAPTER III

METHODOLOGY

Overview

This research describes the differences in the outcome and the process of formative evaluation as carried out by content experts and instructional designers.

The think-aloud method of data collection was used to capture the verbal utterances of subjects during the performance of a formative evaluation task. The data were transcribed verbatim, and coded according to a scheme which was developed on the basis of Human Information Processing Model and terminology. Inter-coder reliability was ascertained by using two independent judges. Comparison between groups was based on means and standard deviations, as well as qualitative characteristics of the coding categories. Several minor coding schemes were also applied, in order to reveal more detailed information for within group comparisons.

Methodological Rationale

Much of the research in the area of formative evaluation has been conducted using experimental or quasi-experimental research designs. Typically the

effectiveness of a revised version of an instructional text has been assessed based on the performance scores of students on criterion referenced tests, or by total time on task (Golas, 1982; Rosen, 1968). Comparisons have also been made between revisions based on comments from learners or those from experts (Frase, de Gracie, & Poston, 1974; Golas, 1983). Statistical techniques, such as analysis of variance and non-parametric tests, have been used to analyze data from these studies and to draw inferences.

Quantitative methods are appropriate when the purpose of an investigation is to demonstrate the effectiveness of formative evaluation or to compare the quality and quantity of input from various conditions, and to draw generalizable conclusions about the superiority of one source over another. However, these methods and the corresponding use of large samples do not provide the level of detail which is necessary to identify the actual processes and activities which take place during the course of formative evaluation. Furthermore, the few studies which have examined the effect of formative evaluation have failed to give proper weight to the actual revision process as a variable. The process of translating the data gathered in formative evaluation into actual revisions has been the focus of a limited number of studies (cf. Dupont & Stolovitch, 1983; Gropper, 1975). The heuristics of instructional design generally suggest that after gathering formative data, the materials be revised "accordingly" without further defining "accordingly". Hence,

a given problem could conceivably be 'fixed' in as many ways as there are revisers. As a result, it is seldom established whether the improvement of the materials was due to the process of formative evaluation and subsequent revision, or due to the skill of a particular reviser. As Kuipers and Kassirer (1984) have pointed out, "individual variation is such a striking feature of human cognition that any attempt to average data across population is certain to mask the true structure of the knowledge" (p. 365).

To study the processes involved in the performance of a task, it has been suggested that detailed descriptive records of the thinking processes of subjects be collected and analyzed (Ericsson & Simon, 1984; Hayes & Flower, 1983). A qualitative approach may be timely in the early stages of research in an applied area when the purpose is to provide a descriptive basis for identifying and defining variables. These variables, in turn, could become the object of manipulation in future, quantitative studies.

As the current study was designed to be an exploratory study, and the task was characterized as ill-defined, it was decided that a detailed record of the performance of three expert subjects in each expert group would yield sufficient raw data to gain insight into the formative evaluation process. This decision was made with the awareness that the criteria which qualify an individual as an expert would have to be defined a priori, and that the selection of subjects would be based upon it.

Selection of Experts

In order to establish the criteria for selecting expert participants, theoretical and empirical frames of reference were sought. Anderson's (1982) theory of skill acquisition describes the stages of skill acquisition which eventually lead to expert status. This theory postulates two phases of knowledge acquisition: an initial declarative phase in which factual knowledge is gained and which is represented in a propositional network of facts; and a procedural phase, where the acquired knowledge is applied in performance.

A review of the empirical literature suggests that the number of years of experience and professional practice in a given domain is the most overt and objective defining characteristic of an 'expert'. For instance, in a study designed to compare procedural differences between expert and novice map readers, Thorndyke and Stasz (1980) included in their expert group a retired officer who had an entire career of map-reading experience, an Air Force pilot who relied on his map-reading skills on a daily basis, and a scientist who was also an amateur cartographer. In addition, all three subjects were teaching map-reading to new recruits at the time the experiment took place. Similarly, Egan & Schwartz (1979) used skilled electronics technicians with over twenty five years of experience working with various types of electronic circuits in their exploration of memory for symbolic circuit drawings. Others,

such as Simon and Simon (1978), and Larkin, McDermott, Simon, and Simon (1980), used expert subjects with strong mathematical backgrounds and wide problem solving experience in their respective studies on the physics problem solving skills of experts and novices.

Expert participants in investigations on writing and revision have also been selected from among professionals. For instance, in a study on the role of revision in the writing processes of skilled and unskilled writers, Sommers (1980) included experienced professional writers such as journalists, editors, and academics. Breuleux (1987b) used professional journalists in his investigation of the planning processes of expert writers. Similarly, Faigley & Witte (1981) utilized the data produced by writers with journalistic and publishing experience in developing their taxonomy for analyzing revision. Beach (1976), and Hayes et al., (1987) used professional editors with teaching experience as judges to rate revised drafts, and experienced writing teachers who had also performed editing tasks, in their endeavor to develop a new model of the revision process in written composition.

Following the precedence set by the cited studies, the criterion established for participation in the current study was that subjects have a minimum of at least eight years of teaching and/ or professional experience in the particular domain (i.e., microbiology or instructional design). Based on this requirement, and after an initial contact to assess

availability, two lists of eight experts in each, Microbiology and Instructional Design, were developed. These lists consisted of eight microbiologists who were teaching Beginning Microbiology to students in the Health Sciences, and professional Instructional Designers whose specialization was the development of instructional materials for various training programs. Three names were randomly selected from each list and designated as the participants of this study.

Method of Data Collection

To investigate performance in a problem solving situation, Newell & Simon (1972) and Faigley, Cherry, Jolliffe, & Skinner (1985) have suggested several data collection techniques. These include concurrent think-alouds, retrospective and introspective interviews, and stimulated cueing.

While all of these techniques can produce rich qualitative data, concurrent verbalizations, or think-alouds, and subsequent analysis of protocols are thought to be particularly appropriate for describing complex problem solving tasks such as revision (Flower & Hayes, 1977; Swarts, Flower, & Hayes, 1984, p.65). As a component process of writing, revision is a complex task because the successful completion of the task cannot be

gauged at the start of the task, nor can it be broken down into well-defined hierarchical steps. Since the evaluation task in this study was aimed at soliciting revision comments from the subjects, it was decided that the think-aloud method would be a suitable means of data collection. In addition to think-alouds, it was decided that retrospective interviews would immediately follow the think-aloud sessions. The purpose of the retrospective interview was to provide subjects with a second chance to elaborate on issues which they felt had not received adequate attention during the think-aloud session. The volume of the data depended upon the quality of comments forwarded by the experts during the formative evaluation task. These data were used only to ease the process of coding the segmented protocols.

In a think-aloud session, subjects are required to verbalize everything that passes through their mind during the performance of a task. (See Appendix III.) The choice of the think aloud method was made with an awareness of some of the criticisms extended toward protocol analysis in general, and think-alouds in particular. For example, Black, Galambos, & Reiser (1984) have asserted that this method interferes with the writing process, resulting in compositions which are inferior in quality and less extensive than if they were written in silence. Similarly, Cooper & Holtzman (1985) have contended that the verbal reports of subjects fail to reveal any of the writing processes which are, by nature, quite complex.

These criticisms, however, seem to be minor considering the intent of this particular study. Based on comparative studies of subjects performing a task in think-aloud versus silent conditions, Ericsson & Simon (1980 and 1984; Hayes & Simon, 1974) concluded that although verbalization may impede the speed of performance, it does not change the structure or course of the cognitive process.

Regarding the strength of this technique in bringing thoughts to the surface, Ericsson & Simon (1980 and 1984) have asserted that a subject's verbalization may not be a completely open window to the internal mental processes, but it can display some of the overall processes, which are utilized at the time of performing a task. This access is expedited by the degree of expertise. An expert can have almost immediate access to domain information that is relevant to the particular task at hand. Ericsson & Simon (1984) have suggested that acquiring expertise involves developing a very systematic way of sorting and storing information. As a result of this unique filing system, when specific knowledge is required to either solve a problem or to answer a question, the relevant information is quickly retrieved.

Having considered these advantages and disadvantages, and given the fact that the experimental task did not require generating new text, nor revising one's own writing, the think-aloud method of data collection was selected for the current study.

Pilot Testing

In order to discover and overcome any potential logistical obstacles and to ensure a smooth data collection procedure, a pilot test was planned and carried out. This pilot study was identical to the projected study in all aspects of the methodology with the exception that the participating subjects were limited to one expert per group. Both subjects had over eight years of professional experience in their respective fields. Data from the pilot study were used to develop the coding scheme for the present study, and were not included in the actual study. This coding scheme is described in the 'Method of Analysis' section.

Procedure

Experimental Task

The experimental task required that subjects rely on their expertise to revise a unit of instructional text on microbiology. Subjects were directed to revise the text to the point where they felt the materials were suitable and ready for publication and use. It was stated that any changes made to the text would be acceptable as revision with the exception of the entire rewriting of the text. Subjects were asked to write their comments on the text

itself, or on an attached blank sheet of paper.

While evaluating the text, subjects were also required to think aloud, verbalizing everything that passed through their minds. All verbal utterances produced during this period were recorded. While the actual performance was not timed, it was noted that the activity took anywhere between one hour and one and one half hour for each individual.

Development of Task Description

The task description was prepared in several stages. It was initially written by the experimenter and revised based on comments from three fellow doctoral students who had had extensive writing and editorial experience with instructional materials. The revised version was then given to two university professors, and two undergraduate students for comment. Subsequent modifications were based on the elicited comments. These included stating explicitly what was meant by 'revision' and elaborating on the think aloud task. (See Appendix I for the Task Description).

Material

The stimulus text used for the experiment was extracted from the draft version of a self-instructional module on Microbiology which had been developed for first year dental hygiene students, "to help prepare the student to become a highly qualified entry-level practitioner" (D.A.E. Project, 1978, pp. xv). The team responsible for the development of

the material comprised content experts, instructional designers, and evaluators. Of six units, only the third titled The Relationship of Microorganisms to Environmental Conditions, Disease and Immunity was used for the study (see Appendix II). (The first two units comprised a Preface and an Introduction, and were not sufficiently specialized to make the content remote for the Instructional Designers.)

The unit was twelve pages long. It consisted of eight pages of instructional text on environmental conditions and the role of microorganisms in producing disease. The text was prefaced with a set of objectives and was followed by a self-test of fourteen 'multiple choice' and 'fill-in-the-blank' questions. Also included was an answer key for the self-test.

The materials were deemed suitable for this experiment for two reasons. First, the specialized content made it feasible to ascertain that experts from either group did not possess knowledge of the each others' domain. Second, the text was in draft form and suffered from numerous shortcomings such as typographical errors, poor spacing and layout, and inconsistency in format. These shortcomings made the reviewer task a realistic one.

Subjects

Six experts were used as subjects in this study. Three were professional instructional designers and three were microbiologists. All six subjects had advanced degrees in

their respective areas. Four of the subjects, two in each of the groups had doctoral degrees, while the remaining two had masters degrees. Aside from academic qualifications, all participants had considerable professional and/or teaching experience in their field: Of the three microbiologists, one was a university professor at the Pathology Department of a Canadian University, who had authored more than sixteen textbooks on this topic. The two others taught beginning microbiology to college students, including dental hygiene students. Each subject's experience in this capacity exceeded the preset criterion of eight years.

Of the three instructional designers, one was currently a university professor, and specialized in developing self-instructional materials for distance education. The second participant had been a university professor prior to forming a consulting company which specialized in the design and development of training materials and programs for business and industry. The third instructional designer had over ten years of experience in developing and evaluating training materials for the private sector.

None of the subjects within either group had specific knowledge about the other group's area of expertise. Each group was composed of two females and one male. Subjects were compensated for their participation in this study.

Experimental Procedures

Preliminary Procedures: Prior to inviting the subjects to participate in the study, the experimenter had conducted a telephone interview in which each potential participant was informed about the topic and purpose of the research, and was given a description of the experimental task. Based on the pilot study, potential participants were also told that the task would require a maximum of two hours.

Upon agreeing to participate, dates were arranged with each participant to carry out the study. At the appointed time, the experimenter met the subject at his/her office. After introductions, the experimenter proceeded by refreshing the participant's memory about the topic and purpose of the research.

Equipment: In order to record the verbalization of subjects during the course of the task, a Sony cassette recording device equipped with a separate microphone was used. The microphone was placed on the subject's desk. The experimenter seated herself outside the subject's immediate viewing range and placed the tape recorder close to herself, where she could monitor the ongoing recording and change the tapes as unobtrusively as possible.

Procedures: Experimental sessions consisted of; a) an orientation to the think-aloud method, b) several think-aloud warm up exercises, c) the formative evaluation of the experimental materials , and d) a retrospective interview. The initial orientation to thinking aloud was

planned because it had been previously established that verbalizing thoughts during a task was a novel activity for all of the subjects. Hayes et al. (1987) have compiled a number of questions that have been brought up by subjects while performing think-alouds. They have provided answers and examples for these questions and have used this text as a set of training materials for thinking aloud. Parts of this document were adapted and given to the subjects to read in preparation for the task. (These instructions can be found in Appendix III.) Following that, a short, recorded think-aloud performance tape, an example of a session, was played for the subject.

Newell and Simon (1972) have also recommended that subjects be acquainted with the procedure of think aloud by practising with several warm-up exercises. One advantage of this activity, among others, is that it allows the subject to become used to talking while thinking. (For a detailed explanation of the merits and logic behind the suggested exercises, the reader is referred to Newell and Simon, 1972, pp. 240-241).

Two such exercises were included in this phase of the study. These involved two questions: one was to name twenty animals, and the other was to report the number of windows in the subject's residence. These exercises provided an opportunity for subjects to become acquainted with the think-aloud method and to hear themselves perform a mental search process.

Upon completion of these two exercises, the subject was given the Task Description sheet and was asked to read it. When finished, the subject was given another chance to ask any questions that remained unresolved regarding the procedures of the task.

At this time, the experimenter reminded the subject that in the event of a pause exceeding ten seconds, she (the experimenter) would prompt the subject to think aloud. It was also reiterated that performance was not being timed, nor was the self-test included in the materials, to be used as a means of evaluating the subject's performance.

The experimenter remained seated out of the immediate view of the subject and followed the subject's progress on her own copy of the stimulus text. Where comments were elaborated inadequately or were postponed pending further reading, the experimenter made a notation on her copy. These notations formed the content of the follow-up retrospective interview. The formative evaluation task was carried out smoothly and was interrupted only if the subject addressed a question to the experimenter. These interruptions, however, were few. Upon completion of the task, the subject's copy of the text and other used sheets of paper were collected.

The retrospective interviews were not pre-structured. As mentioned above, their course was directed by the types of comments which each subject had made during the actual evaluation of the text. Thus their length varied. For instance, if during the task the subject said, "there's a

problem here, but let's see what they do about it later on", and then carried on without returning to the noted problem, during the following retrospective interview, the experimenter probed the subject for further comments. Data from these interviews were only used as corroborative material. The completion of the interview marked the end of the data collection procedure.

At the end, the subject was requested to sign a Participant's Consent Form (see Appendix IV) and was compensated for the time spent on the task.

Methods of Analysis

Transcription and Parsing of Protocols

The recorded verbalizations were transcribed verbatim, using a Sony Transcriber. Sections which were read from the text and the comments which were written, were incorporated into the verbatim transcription (hereafter referred to as a protocol). Transcripts were typed double-spaced, in a continuous format. Pauses were marked by a hyphen (-), and sections which were read verbatim from the module were placed in parentheses, while comments written by the subject were underlined. (See Appendix VI for a sample section from each participant's protocol.)

Typically, transcribed protocols are parsed into segments or units which are then coded for subsequent

analysis. Various methods have been recommended for parsing or segmentation. Speech bursts, temporal information, repetitions, and clauses are some of the suggested ways to segment protocols (Ericsson & Simon, 1980, p. 5, Flower et al., 1980).

Transcripts of protocols from this study were segmented into clausal units since a clause contains sufficient information for making subsequent encoding decisions. The segmentation procedure used in this study was adopted from Winograd's (1972, 1983) system of clausal analysis which is in turn based on Halliday's (1967a; 1967b; 1968) systemic grammar. In this procedure, a tensed verb generally identifies a clausal unit. However, there are some exceptions to this general rule. For instance, a secondary clause can also be considered a separate segment if it is a bound adjunct, that is, if it modifies another clause and is connected to that clause by a binder such as "if", "because", "while", etc. (For a detailed description, the reader is referred to Dillinger, 1987). Each segment was numbered and placed on a separate line.

Protocol Coding

To code parsed protocols, it is essential that a scheme, established a priori, be utilized. This scheme has to be based on the terminology of the theory or model adopted in a given study (Ericsson & Simon, 1980; Swarts et al., 1984). As was discussed earlier, one original aspect

of this investigation was studying formative evaluation within the framework of Human Problem Solving Theory. Of the four major assumptions of this theory, the first suggests that "subject's behaviour [in a problem solving situation] can be viewed as a search through a problem space, accumulating knowledge about the problem situation" (Ericsson & Simon, 1980, p. 263). The elements comprising a problem space include, a) an awareness of a problem element, b) a distinction between the current state of the problem and the goal state, c) a knowledge of how the current state may be transformed to the goal state, and d) knowledge of the scope of this operation. (Hayes and Simon, 1974, p. 167)

In formative evaluation terms, then, the elements which comprise a problem space include a) identifying a problem, b) establishing criteria for an acceptable outcome (i.e., revised product), c) implementing change, and d) drawing upon knowledge sources to achieve the set goal. In translating these elements into coding categories, it was "essential to adhere to elements of the adopted theoretical model (Krippendorff, 1980, p. 75).

The components of the coding scheme which have evolved from the elements of the problem space as well as behavioral and cognitive task analysis of formative evaluation (Johnson, 1988; Ohlsson & Langley, 1985), were initially applied to the pilot data. Subsequent to this application, expansions and refinements were made. This scheme is

presented in the next section.

Major Coding Scheme

The major coding scheme of this study was applied to all the parsed segments. The entire system consists of 13 categories which are first described and then followed by examples from subjects' protocols.

Text Related Categories

1) Evaluation Statements:

The outcome of a comparison between the currently observed state and the goal state, that is what the reviser thought the text ought to be, was coded as an Evaluation Statement (ES). By virtue of this definition, this category represented positive and negative comments (which did not explicitly state the source of the problem), expressions of preference, judgement, internal feelings and observations, all expressed in the context of the subject's task representation.

- *I don't particularly like it/*
- *you've got a lot of white space here which is fine/*

2) Problem Identification:

As a specific case of Evaluation Statements, Problem Identifications (PI) contained explicit references to an observed problem.

- *again, I find the heading's really not well done/*
- *and right justified margin makes the question difficult to read/*

3) Revision Statements:

Explicit text related changes intended to transform the current state to the goal state were coded as Revision Statements (RS).

- *there should be a comma after 'disinfection'/*
- *'cross infection' should be hyphenated/*

4) Knowledge Statements:

Expressions of personal knowledge, including both declarative and procedural knowledge, were coded as Knowledge Statements (KS). These statements were often provided in addition to, or in lieu of, a Problem Identification, or as a reason for suggesting a particular change.

- *I don't particularly like it/ (ES)*
- *most microorganisms especially those that cause disease don't require light at all/*
- *and the orthotropes which do require light are not important in medical microbiology/*

- *that's true only to a certain extent/ (ES)*
- *but more importantly would be the virulence of the micro-organism, how severe, how effective is it in causing disease/*
- *so it is more important to consider the property rather than the number of micro-organisms/*

5) Text Knowledge:

Comments representing Knowledge directly acquired from the stimulus text were coded as Text Knowledge (TK).

- *um, so I can assume that cross-infection is not from your body/*
- *that's one type of chemical, but that's already in* the stomach/

6) Verbatim Statements:

Sections which were read aloud from the text were coded as Verbatim Statements (VS).

7) Text Talk:

Segments which referred to various parts of the text, but not in verbatim format, were coded as Text Talk (TX).

- *so there's a self test to test these objectives/*
- *so its gonna go into talking about those three/*

Task Related Categories

8) Task Talk:

Reference to an activity which was currently being undertaken or to the set up of a short term goal was coded as Task Talk (TT).

- *I'm reading the introduction/*
- *let's see if they discuss this in the next section/*

9) Strategy Talk:

Reference to a course of action which was representative of actions normally undertaken by the subject in similar situations, but were not tied to the current task were coded as Strategy Talk (ST).

- *I usually check to see if the objectives match the text/*
- *I like to get a sense of the material first, so I'll* just skim through/

Four categories captured the remaining segments. These included Dialogue (D) which encompassed questions addressed

to the experimenter, and the experimenter's subsequent response; Boundary Markers (BM) (e.g., ok; um; uh); False Starts (FS); and Unrelated Talk (UT). The above mentioned thirteen categories allowed all segments of the protocols to receive mutually exclusive coding as has been recommended by Krippendorff (1980). However, the last four (D, BM, FS, BM) were neither theoretically significant nor did they provide interesting information. Thus, they were excluded from the analysis.

The application of this system to the verbal data rendered a manageable database that could be submitted to more detailed levels of analysis.

Coding Reliability

In qualitative research, the reliability of a system which is used to code, translate and interpret data must be ascertained before one can proceed to a meaningful discussion of results. The two types of reliability, reproducibility and stability, which have been identified as relevant to content analysis (Krippendorff, 1980; Weber, 1985), were adhered to in this study. Reproducibility or inter-coder reliability was established by assigning two independent judges, one of whom was the experimenter, to perform the coding of all segmented protocols. The coders used a coding sheet (see Appendix VI) which described categories in the scheme and provided a range of typical examples. To acquaint the second coder with the system, data

from the pilot study (which were not included in the analyses) were used as practice material. One protocol was coded by the second judge with the assistance of the experimenter. During this practice period, ambiguities regarding definitions were clarified and questions were answered. The second coder was then assigned the task of independently coding the protocols of the six subjects, relying only on the coding sheet. It was suggested that segments be double coded only if their exclusivity to one category could not be firmly established. The coder was compensated for performing the task.

Protocols coded by the two judges were then compared. Overall inter-rater reliability was 89%. Reliability on Problem Identification, Evaluation, and Revision Statements was 88%, 94%, and 95%, respectively. Segments which received double coding amounted to less than 2%, and did not include Evaluation, Problem Identification, or Revision Statements.

The second type of reliability, stability, that is the invariance in coding the text over time was established by the experimenter who coded all protocols three times with one week and one month time lapses for each recoding, respectively. This procedure was carried out with 89% reliability the second time, and 87% the third time.

Frequencies, Percentage Scores, and Descriptive Statistics

After ascertaining reliability in coding, a database was created by tabulating the frequency of occurrence of each category per expert. Frequencies were then converted into percentage scores, based on the total number of segments per protocol. A second set of percentage scores were calculated, excluding theoretically irrelevant categories, and repetitions. The irrelevant categories were comprised of Boundary Markers (BM), False Starts (FS), Unrelated Talk (UT), and Dialogue (D). The first and last segments in the example below represents a repetition. In this case both were coded as one Revision Statement as they implied the same revision:

- *I would eh definitely rewrite this/*
- *eh/*
- *its a bit confusing/*
- *it doesn't say enough/*
- *its too much and not enough at the same time/*
- *so I would expand it/*
- *I would rewrite/*

However, when a statement referred to an independent aspect of change, it was tabulated separately. In the above example, the two last segments suggest two types of revisions: One is the addition of content, and the other is a change in wording. These two were coded as two separate Revision Statements.

Descriptive statistics was applied to this database to compare between groups means.

Minor Coding Schemes

Classification of Evaluation, Problem Identification, and Knowledge Statements.

While quantifying data did provide information regarding amounts, delineating the type of comments that were being prompted by the experts' particular expertise, depended upon a detailed content analysis of protocols. A secondary coding scheme was applied to the four major text-related categories of Evaluation, Problem Identification, Revision, and Knowledge Statements. Within each of these four categories, comments were classified as being related to Content, Design, Presentation, and Pedagogy. This implied that comments which made reference to the subject matter, were classified as Content, those which referred to instructional design heuristics were coded as Design, those which dealt with the physical appearance of the text were categorized as Presentation, and those evaluating the text in light of a potential learner were considered to be Pedagogical. These areas reflected the range of expertise of the participants.

Since the reliability of this coding scheme had been established in a previous study (Saroyan and Geis, 1988), it was not deemed necessary to duplicate the effort.

Descriptive statistics were applied to the data in order to create a basis for between group comparisons.

Classification of Problem Identification and Revision Statements.

In order to derive information on Problem Identification and Revision Statements on a deeper level, these statements were analyzed according to a) the degree of specificity, and b) convergence and divergence.

A: Degree of Specificity: The purpose of this analysis was to delineate the specific features of text which drew comments from the subjects. It further allowed highlighting comments which were made more frequently, and those which were rare and idiosyncratic. Two comprehensive listings of all comments generated by each subject regarding Problem Identification and Revision were created. Percentage frequencies of these comments were then tabulated in order to perform within group comparisons.

B: Convergence and Divergence: The purpose of this analysis was to determine the degree of convergence and divergence between and within groups on Problem Identification and Revision. A database was created by transferring onto a master copy of the stimulus text Problem Identification and Revision Statements of the six subjects. (This text was segmented into naturally occurring sentences. See Appendix II.) The outcome indicated a) segments which had received mutual Problem Identification and/or Revision Statements; b) segments which had received mutual Problem Identification but different Revision recommendation, and c) segments which had received different Problem Identification comments, but similar Revision recommendation.

Classification of Revision Statements.

Revision Statements were categorized according to two additional systems: a) by type and level of change, and b) by their components parts.

A: Type and Level of Change: To establish a profile of the type of change, and to discern whether the revision suggestions were made on a local level, like novices, or a global level, like experts, an existing classification system was utilized. This system, which is often referred to in the writing literature, is used to classify revision changes of expert and novice revisers (Bridwell, 1980; Faigley and Witte, 1981; Sommers, 1980) by addition, deletion, substitution, and rearrangement, and by the level at which change occurs.

B: Revision Component: To determine the degree of elaboration of suggested changes, Revision Statements were analyzed in terms of their constituent parts. Richard's (1986) proposed framework on "Knowledge of Action", was used as model for designating the components of a Revision Statement.

The components of a Revision Statement were identified as an action, a result, and a rationale. To operationalize this analysis, the following definitions were developed: **Revision Action** is marked by a non-specific revision operator such as 'change', 'improve', or 'clarify'. This statement suggests a change, but does not specify any details about implementing the change. Hence, in order to

arrive at the desired goal state, it is conceivable that several acceptable courses of action may be taken. For instance to implement a change one could add, delete, substitute or reorder content, or presentation. **Revision Result**, on the other hand, is marked by a specific revision operator which limits the means of implementing the change to one particular method. This includes those revision suggestions which have been conveyed in writing as well those which have actually been spelled out or recommended verbally. For instance, *"this should really be in Fahrenheit rather than Centigrade"/* conveys a specific change. **Revision Rationale** refers to the reason for suggesting the change. This reason, in turn suggests whether the motive for change is individual, that is based on personal like or dislike, or conventional and based on a particular model or heuristic.

In the example below, the last segment is the rationale:

- *this would not destroy endospores/*
- *and that should be emphasized here/*
- *without it, it is incomplete/*

Percentage scores across subjects were calculated for both sets of analyses, and comprised the database for between and within group comparisons.

Summary

The most significant aspect of the methodology of the current study was the development of a theoretically grounded coding scheme which could be effectively applied to verbal data generated during formative evaluation. Subsequent quantitative and qualitative analysis of the coded protocols provided a basis for between and within groups comparisons.

CHAPTER IV

RESULTS AND DISCUSSION

Overview

In a descriptive, exploratory study, general questions are posed at the outset to guide the research. However, data often produce results which are quite beyond the scope of the posed research questions. The significance of such findings, in turn, is that they contribute to future theory building and hypothesis generating and testing.

The purpose of this study was to delineate the processes which content experts and instructional designers engage in during formative evaluation. Four direction giving research questions were posed at the outset of this study. These questions required a detailed recording of the performance of content and instructional design experts in formative evaluation of an instructional text, and the subsequent comparison between and within group members. Specifically, it necessitated that a) the elements comprising the problem space such as evaluation, problem identification, revision, and the processes and strategies used in formative evaluation be quantified and described, and b) based on the above data, comparisons be made between and within the expert groups.

This chapter comprises results pertaining to the

research questions followed by the ensuing discussion. The section concludes with a general discussion of the findings.

Between and Within Groups Results and Preliminary Discussion

The results presented in this section include frequency data on the categories, between groups mean data, within group percentages, and proportions of distribution of comments on Content, Design, Presentation, and Pedagogy.

Mean Frequencies of Coded Categories

As a first step, a database of frequencies of coded segments per subject protocol was created, and converted into percentage scores. (See Tables 1 and 2 in Appendix VII-A and VII-B.) A second set of percentage scores was tabulated by excluding repetitive statements and the categories which were theoretically irrelevant. (The irrelevant categories included Boundary Markers, False Starts, Unrelated Talk, and Dialogue.) Percentage of frequencies is displayed in Table 3).

The nine categories presented in this table, were either text or task related. Of the text related statements, Evaluation, Problem Identification, Revision, and Knowledge Statements were of primary importance because they comprised phenomena within the problem space which signified the outcome of formative evaluation. The

remaining three, Verbatim Statements, Text Talk and Text Knowledge were of secondary importance statistically, but significant qualitatively since they depicted the way in which the text was used by each subject. Task related statements included Task Talk, and Strategy Talk, and represented the particular strategic processes of each subject.

Table 3

Percentage of Significant Categories by Expert

	Content Expert			Instructional Designer		
	I	II	III	I	II	III
Evaluation	25.4	11.9	20.4	18.0	17.7	29.5
Problem Id.	5.9	9.3	7.6	11.5	17.7	12.3
Revision	12.5	11.3	10.4	6.4	9.5	12.0
Knowledge	12.2	31.0	12.4	2.3	11.3	4.0
Text Know.	2.0	0.0	1.7	7.8	6.3	2.1
Verbatim	29.0	27.4	21.1	38.8	18.6	9.3
Text Talk	5.9	2.5	7.6	8.3	9.3	5.5
Task Talk	4.4	2.5	5.8	3.7	5.0	14.9
Strategy Talk	0.0	0.5	0.0	1.8	2.2	1.7
TOTAL (# of segments)	334	193	458	216	439	794

Figures 2 and 3 display the percentage of the frequency of occurrence of these nine categories by expert group.

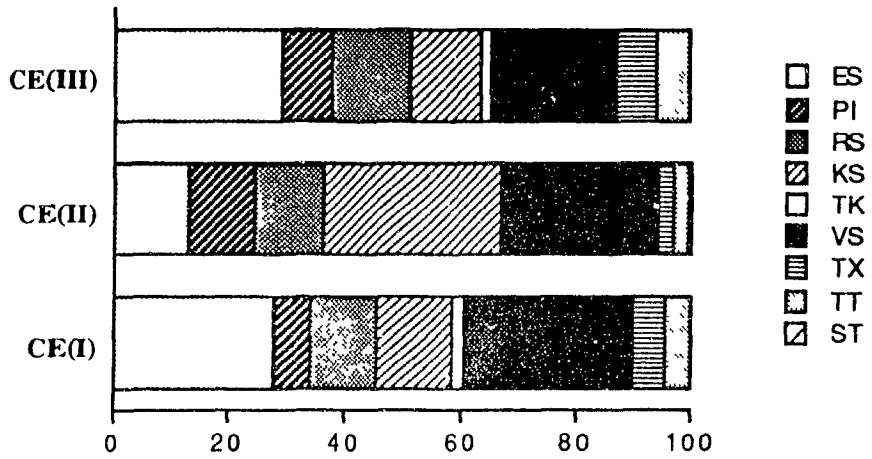

FIGURE 2.

Percentage of categories per content expert

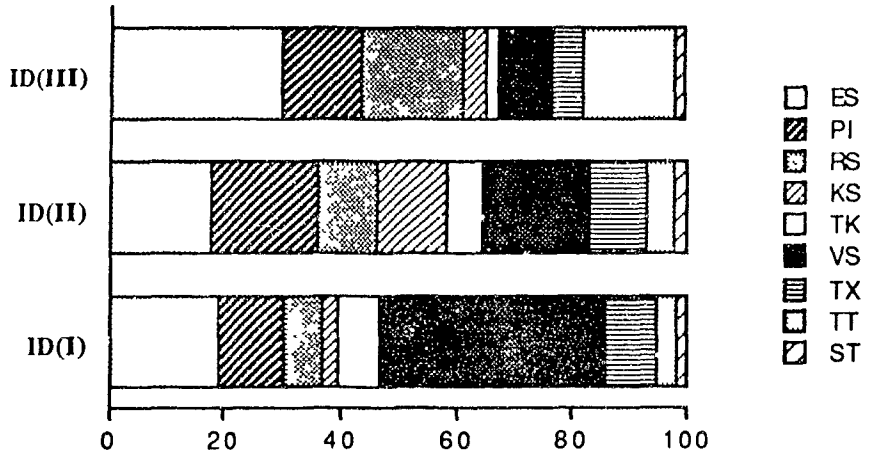

FIGURE 3.

Percentage of categories per instructional designer

The remainder of this section is organized around between groups means, and within groups percentage scores by the categories in the initial coding scheme. The first table presented for each category depicts between groups mean data. In the upper half of each table, mean group frequencies, and percentages, (and standard deviations) are presented. In the bottom half of the table, where relevant, the distribution of Statements among the four subcategories; content, design, presentation, and pedagogy, is shown as group mean percentages. Where relevant, more detailed data on idiosyncratic behavior of each subject are provided.

Text Related Statements

Evaluation.

All expressions of internal feelings, including judgement, opinion, and observations about the stimulus text conveyed within the context of the formative evaluation task, were coded as Evaluation Statements. Group results are displayed in Table 4.

Table 4

Between Groups Mean Data on Evaluation Statements

	Content Experts	Instructional Designers
$\bar{X}$ Frequency	76.6	117.3
$\bar{X}$ Percentage (S.D.)	19.2 (6.8)	21.7 (6.7)
$\bar{X}$ % Content	97.1	30.4
$\bar{X}$ % Design	0.3	39.2
$\bar{X}$ % Presentation	0.0	14.0
$\bar{X}$ % Pedagogy	2.4	16.2

The mean percentages of the two groups indicate that Evaluation Statements comprised about one fifth of the protocols. Between group disparity was more prominent in the proportions of evaluations statements representing the four categories of Content, Design, Presentation, and Pedagogy. The interesting finding was that nearly all Content Experts evaluated only Content (i.e., the factual information in the text), while the designers also evaluated other aspects including design, presentation, and pedagogy.

More significant than the quantitative differences were the qualitative differences in the statements of the two groups. While the Content Experts' Evaluation Statements represented either approval or disapproval of the factual information provided in the text, or an observation of an existing situation (Figure 4), the Instructional Designers' also represented the degree to which the text adhered to certain design criteria.

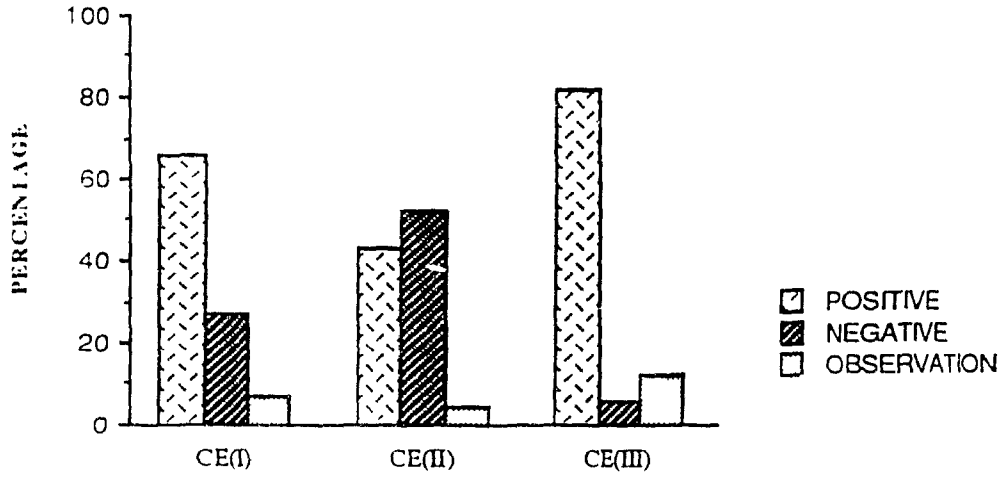

FIGURE 4. Type of Evaluation by content expert

The criteria or heuristics referred to in the evaluation of all three Instructional Designer included:

1. Alignment between various parts of the text:
   a. objectives with text
   b. objectives with test items
   c. test items with text
2. Validity and quality of objectives.
3. Logical sequencing of text.
4. Job-relatedness of content and examples.
5. Matching level of content with entry behaviour.
6. Adherence to presentation principles such as:
   a. use of visuals
   b. use of headings
7. Pedagogical value of:
   a. content
   b. test items
   c. the curriculum in general

Additionally, the Instructional Designers partially based their evaluation on the success of the text in preparing the naive learner, in this case the subject, for the subsequent self-test.

Within group patterns appeared to be fairly consistent among subjects. Tables 5 and 6 depict the individual data by expertise.

Table 5

Percentage of Type of Evaluation Statement by Content Expert

	Content Experts		
	I	II	III
Content	98.8	100.0	92.6
Design	1.1	0.0	0.0
Presentation	0.0	0.0	0.0
Pedagogy	0.0	0.0	7.3

Table 6

Percentage of Type of Evaluation Statement by Instructional Designer

	Instructional Designers		
	I	II	III
Content	7.6	56.4	27.2
Design	46.1	24.3	47.2
Presentation	17.9	6.4	17.8
Pedagogy	28.2	12.8	7.6

The contrast between the two groups and the similarities within group members is apparent from the results presented in the above two tables. The evaluative comments of all the Content Experts were overwhelmingly Content related, while all the Designers referred to the four aspects of the text. Design issues appeared to be strongly favoured, in particular by Instructional Designer

(I) and (III).

It is important to elaborate on the evaluations which were based on conventional criteria, because it suggests that contrary to some assertions (Baker, 1970; Deisler & McNeil, 1960; Gropper, Lumsdaine, & Shipman, 1961; McEntie & Rivers, 1971), formative evaluation is not entirely based on intuition. The examples which are provided below, have been extracted from the instructional designers' protocols, and they reflect the strong presence of several heuristics in each subject's frame of reference.

Heuristics referred to by Instructional Designer (I) included:

1) Alignment between objectives and the text.

- *first section must have to do with the second objective/*
- *we're still on the same section that seems to subscribe to objective three/*

2) Alignment between test items and objectives.

- *there's a self test to test the objectives/*
- *questions are within keeping of the kinds of things that objectives call for/*

3) Alignment between test items and text.

- *questions are in the same order as the text/*
- *I assume this question could be answered when*
- *reading 'Oxygen and Light'/*

4) Pedagogical value of the text.

- *people are not being asked to manipulate ideas but simply recall/*
- *if someone asked me tomorrow, I think I would only remember a few things/*

5) Arrangement and Presentation.

- *everything is of equal importance supposedly, visually, by inference/*

Heuristics referred to by the Instructional Designer (II) included:

1) The quality of objectives (both unit objectives and the objectives of entire module).

- *I'd say people would have difficulty getting from here to there/*
- *the objectives listed here are relatively ok/*

2) Alignment between objectives and text.

- *diagram of immunity system which relates back to the objective/*

3) Alignment between objectives and self-test.

- *I guess the thing is that you've got objectives here and you have a post-test that does seem to be getting at objectives/*

4) Significance of visuals.

- *they [the students] may need prompts, cues, whatever you put into it to make it more instructional/*
- *we haven't at all talked about audio-visuals/*

5) Match between content and the entry behaviour.

- *I'm not sure if given students' entry behavior, their capabilities and things like that, that they'd be able to do that/*
- *I guess they are assuming some basic knowledge here/*

6) Job-relatedness of content and examples.

- *what we have here is content but not necessarily related to specifics/*

7) Pedagogical value.

- *I wouldn't know if they'd know this stuff by looking at the exam/*
- *I understand what they're talking about but just provide ease of access to the learner/*

Heuristics referred to by Instructional Designer (III) included:

1) Stating objectives in behavioural terms by using action verbs.

- *I like the way the objectives have action verbs/*

2) Gradual progression of objectives from simple to complex, and from cognitive to applied.

- *ok, the objectives seem to move from sort of simple to complex in their sequence/*
- *they seem to go from more cognitive things to more application orientations/*

3) Alignment of objectives with the text.

- *so here is where it [resident microorganisms] comes up/*
- *so going back to my list, cross infection was covered/*

4) Alignment of objectives with test items.

- *they are really not being asked to define it/*
- *they're being asked to recognize a definition of/*
- *I'm looking for its construction as related to the objectives/*

5) Job-relatedness of content.

- *I'm thinking, gee, how does this relate to my world/*
- *I should be concerned about instruments that may have touched one patient and may be touching another/*
- *I should be concerned with checking the temperature on my sterilization equipment/*
- *but what do I care about bugs/*

6) Consistency in format.

- *now this format again/*
- *come on, let's be consistent here/*

7) Pedagogical value of the material.

- *I'm suspecting that if this is for a new dental student, this may be a little bit threatening for* them/

As mentioned earlier, assuming a 'naive learner's ' position and considering performance on the test as a measure of effectiveness of the material was another distinct quality displayed by all three group members. Some examples from the protocols are presented below.

- *and I'm in a panic cause I don't know whether I have learned this information/*
- *so now I know the difference between resident and transient microorganisms/*
- *I understand what they are talking about/*
- *I have no idea what it would mean/*
- *so I find that I was able to complete this exam with only 1 out of 14 without reading the text/*

In summary, while the quantitative profile of the two groups regarding Evaluation Statements resemble one another, the qualitative attributes were strikingly different. The Evaluations of the Instructional Designers' extended beyond the limited scope of the Content Experts' in that they were more global, were prompted by a set of design heuristics, and were partially formed by approaching the text as a naive

learner. Within group results suggested consistency among both group members.

Problem Identification Statements.

Statements which explicitly described an inadequate or unacceptable feature of the text were coded as Problem Identification Statements.

Group data on the Problem Identification category are displayed in Table 7.

Table 7

Between Groups Mean Data on Problem Identification Statements

	Content Experts	Instructional Designers
$\bar{X}$ Frequency	24.3	67.0
$\bar{X}$ Percentage (S.D.)	7.6 (1.7)	13.8 (3.3)
$\bar{X}$ % Content	96.2	32.0
$\bar{X}$ % Design	0.0	27.9
$\bar{X}$ % Presentation	1.8	23.9
$\bar{X}$ % Pedagogy	1.9	15.8

Two important results on group differences are revealed in Table 7. First, the mean percentage of Problem Identification Statements was higher for Instructional Designers than the Content Experts ($\bar{X}$=13.8% > $\bar{X}$=7.6%). Second, the concentration of comments by the Content Experts was on Content issues, while those of the Instructional

Designers related to all of the four possible categories. A particularly interesting finding in this area was the minimal number of pedagogical comments by the Content Experts. Such a finding is significant in light of the fact that each member of the Content Expert group had extensive experience in teaching the subject and was highly qualified to comment on the pedagogical aspect of the material. Interestingly, it was the Instructional Designers who generated comments which were more demonstrative of a concern for the potential user. ($\bar{X}$= 15.8% versus $\bar{X}$=1.9).

The qualitative difference observed in the Problem Identification Statements suggests that each type of expert may have assigned different meaning to their role as a reviewer. The two Task Description sheets given to each group differed in one aspect only, which indeed may have been crucial in the role interpretation. One group was addressed as "Content Specialist ... knowledgeable about dental hygiene students", and the other as "Instructional Designer[s]" . It appear that the Content experts viewed their role as a 'specialist', and the Instructional Designers as a 'generalist'.

Within group variability on Problem Identification appeared to be greater among the Instructional Designers than the Content Experts ($\bar{X}$=7.6%, S.D. 1.7 versus $\bar{X}$=13.8%, S.D. 3.3). Tables 8 and 9 display the proportion of types of problems identified by each of the members of the two groups.

Table 8

Percentage of Type of Problem Identification by Content Expert

	Content Experts		
	I	II	III
Content	100.0	100.0	88.7
Design	0.0	0.0	0.0
Presentation	0.0	0.0	5.6
Pedagogy	0.0	0.0	5.7

Table 9

Percentage of Type of Problem Identification by Instructional Designer

	Instructional Designers		
	I	II	III
Content	4.0	56.3	35.7
Design	28.0	29.2	26.5
Presentation	40.0	2.4	29.5
Pedagogy	28.0	11.4	8.1

Of interest was the repetition of a definite pattern of concentration on Content matters by all the Content Experts (Min.=88.7, Max.=100.0), while the percentage of comments of the designers indicated a distribution among all four areas. Nonetheless, only comments regarding Design were within close range (Min.=29.2, Max. 26.5).

The difference among subjects in both groups became more acute when Problem Identification comments within each category were analyzed and compared on a more detailed level. Table 10 depicts the breakdown of Problem Identification Statements by category and by specific comments.

Table 10

Percentage of Specific Problems by Expert

	Content Expert			Instructional Designer		
	I	II	III	I	II	III
CONTENT						
.spelling	5.0	16.6	2.8	0.0	0.0	5.1
.adequacy	40.0	11.1	31.4	4.0	30.7	10.2
.accuracy	0.0	44.4	2.8	0.0	3.8	0.0
.quality of examples	0.0	5.5	2.8	0.0	0.0	0.0
.redundancy	10.0	5.5	5.7	0.0	0.0	2.0
.terms	0.0	16.6	5.7	0.0	0.0	0.0
.specificity	25.0	0.0	0.0	0.0	0.0	0.0
.consistency	5.0	0.0	5.7	0.0	1.2	1.0
.clarity	15.0	0.0	20.0	0.0	12.8	9.1
.sequence	0.0	0.0	2.8	0.0	2.5	1.0
.relevance	0.0	0.0	8.7	0.0	5.1	7.1
DESIGN						
.objectives	0.0	0.0	0.0	0.0	8.9	7.1
.alignment	0.0	0.0	0.0	8.0	11.5	11.2
.poor design	0.0	0.0	0.0	0.0	2.5	1.0
.inadequate practice/examples	0.0	0.0	0.0	0.0	3.8	4.0
.needs asses.	0.0	0.0	0.0	0.0	2.5	0.0
.link	0.0	0.0	0.0	4.0	0.0	0.0
.organization	0.0	0.0	0.0	12.0	0.0	2.0
.link with previous knowledge	0.0	0.0	0.0	4.0	0.0	0.0
.no rationale	0.0	0.0	0.0	0.0	0.0	1.0
PRESENTATION						
.visuals	0.0	0.0	0.0	0.0	1.2	0.0
.headings	0.0	0.0	0.0	20.0	0.0	0.0
.format	0.0	0.0	0.0	8.0	1.2	13.2
.space	0.0	0.0	2.8	8.0	0.0	0.0
.margins	0.0	0.0	0.0	4.0	0.0	3.0
.typeface	0.0	0.0	2.8	0.0	0.0	5.1
.consistency	0.0	0.0	0.0	0.0	0.0	7.1
.cues	0.0	0.0	0.0	0.0	0.0	1.0
PEDAGOGY						
.retention	0.0	0.0	5.7	12.0	1.2	0.0
.instruct. quality/validity	0.0	0.0	0.0	16.0	10.2	8.1
TOTAL	20	18	35	25	78	98

The advantage of this type of detailed analysis is that it depicts the specific features of text which the subjects render opinion on. As the results indicate, among the Content Experts only three topics (i.e., spelling, inadequacy of content, and redundancy) were periodically referred to by all three group members. Five aspects were mentioned by two, while the remaining problems were pointed out by only one subject.

Among the Instructional Designers, only five problems including adequacy of content, alignment of parts, formatting, and the instructional validity of the material, were raised by all three subjects. Eleven issues were referred to by two, and another eleven by only one of the subjects. It must be reiterated that the identified problems did not necessarily refer to the same source. They merely represented the type of issues which were raised by these two types of experts.

In summary, there was a greater degree of similarity within the experts groups than between the groups (Figure 5).

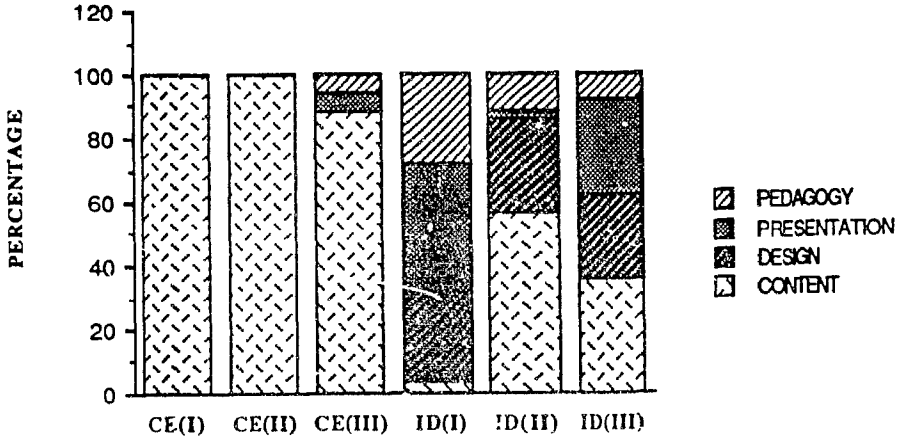

FIGURE 5. **Type of Problem Identification by expert**

The Content Experts appeared to associate problems with the content, in particular its adequacy, accuracy, clarity of expression, amount of detail, and correct spelling. The Instructional Designers, on the other hand, appeared to find problems with not only the design, but also the content, presentation, and the pedagogical value of the text. Consistency in performance patterns of both groups diminished when Problem Identification was compared at a deeper level. Figures 6 and 7 display the distinction both between and within groups.

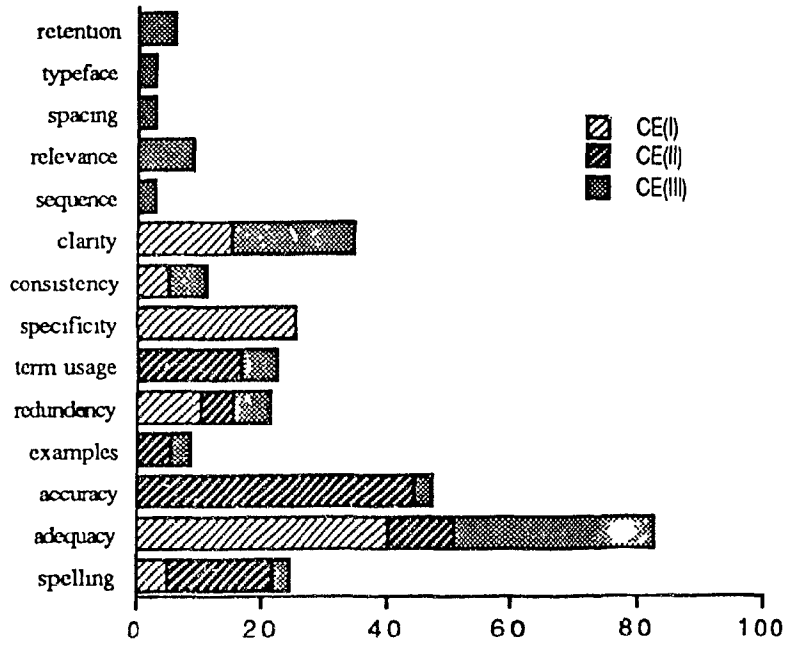

FIGURE 6.

Percentage of Problem Identification by content expert

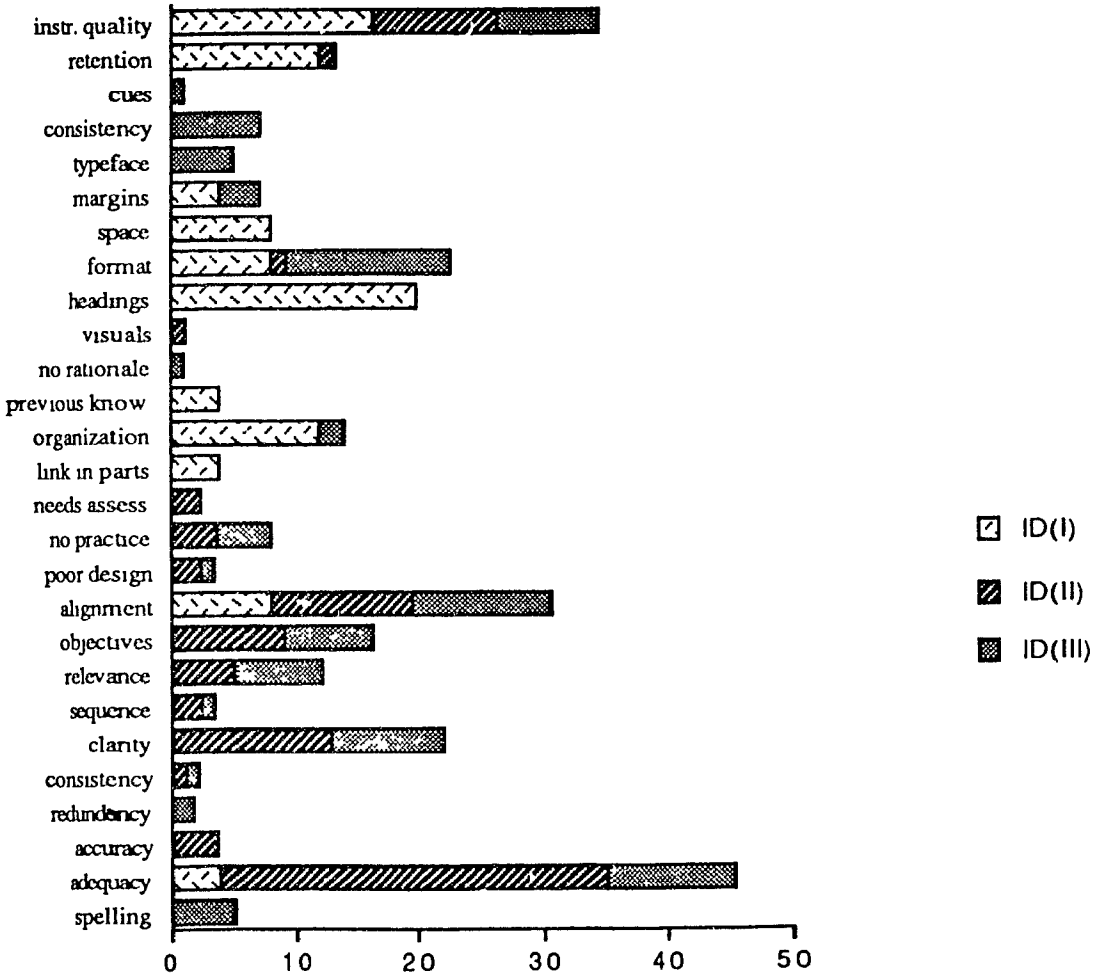

FIGURE 7.

Percentage of Problem Identification by instructional designer

Revision Statements.

A Revision Statement constituted a preferable alternative to a particular aspect of the current text. Overall group data on Revision Statements and percentage distribution by category are presented in Table 11.

Table 11

Between Groups Mean Data on Revision Statements

	Content Experts	Instructional Designers
$\bar{X}$ Frequency	37.3	52.6
$\bar{X}$ Percentage (S.D.)	11.4 (1.0)	9.3 (2.8)
$\bar{X}$ % Content	92.8	28.7
$\bar{X}$ % Design	0.0	50.0
$\bar{X}$ % Presentation	7.0	20.7

As in the previous two categories, the mean percentage of Revision Statements was higher for the Content Experts than the Designers ($\bar{X}$=11.4>$\bar{X}$=9.3). Distribution of recommendations among the four categories resembled Problem Identification in that the majority of suggestions by the Content Experts related to Content, while the recommendations of the Designers applied to Content as well as Design and Presentation. Recommendation on pedagogical improvements were absent from both group members' protocols. This pattern was fairly consistent with the Problem Identification statements of the Content Experts. The absence of such recommendations in the Instructional

Designers' protocols may have been due to the fact that these experts simply did not know how to remedy some of the problems which had been identified by them. The duplication of the 'specialist' and 'generalist' role lends further support to the assertion that the reviewer role was interpreted differently by the members of each group.

Within group variability on Revision Statements was greater among the Designers than the Content Experts. ($\bar{X}$=11.4, S.D.=1.0 versus $\bar{X}$=9.3, S.D.=2.8). Tables 12 and 13 display information regarding the individual proportions of comments by the two expert groups.

Table 12

Percentage of Type of Revision by Content Expert

	Content Experts		
	I	II	III
Content	95.2	100.0	83.3
Design	0.0	0.0	0.0
Presentation	4.7	0.0	16.5

Table 13

Percentage of Type of Revision by Instructional Designer

	Instructional Designers		
	I	II	III
Content	0.0	59.2	26.9
Design	85.6	21.4	43.1
Presentation	14.2	19.0	28.9

These results indicate that the majority of recommendations of all three Content Experts were in the Content area, although two out of three subjects also had a small percentage of comments on Presentation. The recommendations of the Instructional Designers, referred to all three aspects, except Instructional Designer (I) whose recommendations were not related to Content at all. The range was smallest on Presentation (Min. 14.2, Max. 28.9), and greatest on Design (Min. 21.4, Max. 85.6). Nonetheless, the mean percent of Revision comments was greatest on Design matters ($\bar{X}$=50.3 versus $\bar{X}$=28.7 and $\bar{X}$=10.7 for Content and Presentation, respectively.)

When the Revision Statements of the subjects were compared on a deeper level in order to identify those recommendations which appeared to be favoured by each

subject, there was less consistency and more idiosyncratic tendencies across all the subjects. These data are presented in Table 14.

Results indicate that only two specific recommendations, spelling, and the adequacy of content were shared by all three Content Experts. Six were mentioned by two of the three subjects, and five, by only one subject. Consistency among the instructional designers was equally minimal. Only two recommendations were shared by all three Designers (i.e., adding a meaningful structure, and a summary), while eleven were made by two, and twelve by only one subject.

Table 14

Percentage of Specific Revisions by Expert

	Content Expert			Instructional Designer		
	I	II	III	I	II	III
CONTENT						
.spelling	2.3	27.2	2.0	0.0	0.0	5.2
.punctuation	0.0	13.6	2.0	0.0	0.0	1.0
.tense	0.0	4.5	0.0	0.0	0.0	0.0
.adequacy	40.4	13.6	9.5	0.0	11.9	13.5
.relevance	28.5	36.3	0.0	0.0	28.5	0.0
.redundancy	0.0	4.5	4.1	0.0	0.0	0.0
.clarity	7.1	0.0	4.5	0.0	14.2	4.1
.sequencing	14.2	0.0	2.0	0.0	2.3	3.1
.quality of examples	2.3	0.0	8.7	0.0	2.3	0.0
DESIGN						
.objectives	0.0	0.0	0.0	0.0	0.0	7.2
.alignment	0.0	0.0	0.0	7.1	0.0	5.2
.addition of self test/practice	0.0	0.0	0.0	0.0	7.1	11.4
.addition of summary/overview	0.0	0.0	0.0	14.2	4.7	1.0
.addition of examples	0.0	0.0	0.0	0.0	0.0	4.1
.logical dev.	0.0	0.0	0.0	21.4	4.7	0.0
.hierarchical organization	0.0	0.0	0.0	14.2	0.0	4.1
.proximity of text with diagrams	0.0	0.0	0.0	14.2	0.0	2.0
.meaningful structure	0.0	0.0	0.0	14.2	2.3	4.1
.bite-size presentation	0.0	0.0	0.0	0.0	2.3	0.0
.rationale	0.0	0.0	0.0	0.0	0.0	2.0
.consistency	0.0	0.0	0.0	0.0	0.0	1.0
.job-related	0.0	0.0	0.0	0.0	0.0	1.0
PRESENTATION						
.visuals	0.0	0.0	4.1	0.0	11.1	0.0
.headings	0.0	0.0	0.0	7.1	0.0	2.0
.format/layout	0.0	0.0	0.0	0.0	2.3	6.2
.space	0.0	0.0	10.4	7.1	0.0	3.1
.highlighting	4.7	0.0	0.0	0.0	4.7	14.5
.print density	0.0	0.0	2.0	0.0	0.0	3.1
TOTAL	42	22	48	14	42	96

In summary, results on Revision Statements suggest that the Content Experts and the Instructional Designers made different types of recommendations for change. Revisions were mostly limited to Content for the former group and more general for the Instructional Designers (Figure 8).

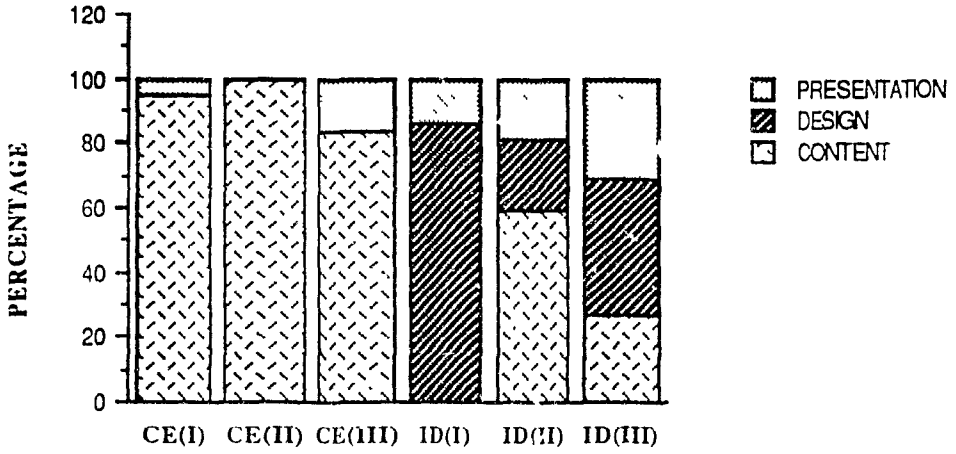

FIGURE 8. **Type of Revision Statement by expert**

A repeated pattern of results among the subjects in each group, lend further support to the 'specialist' role of the Content Expert, and the 'generalist' role of the Designers. Consistency was greatly reduced when specific change recommendations were compared across subjects. This information is summarized in Figures 9 and 10.

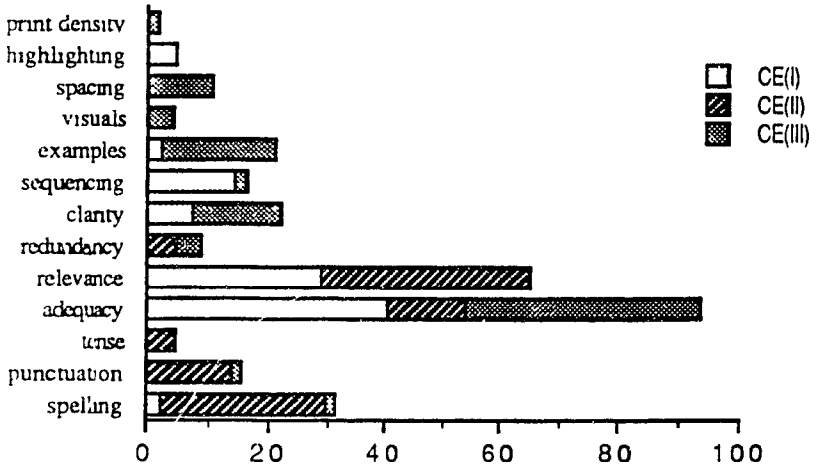

FIGURE 9. **Percentage of type of revision per content expert**

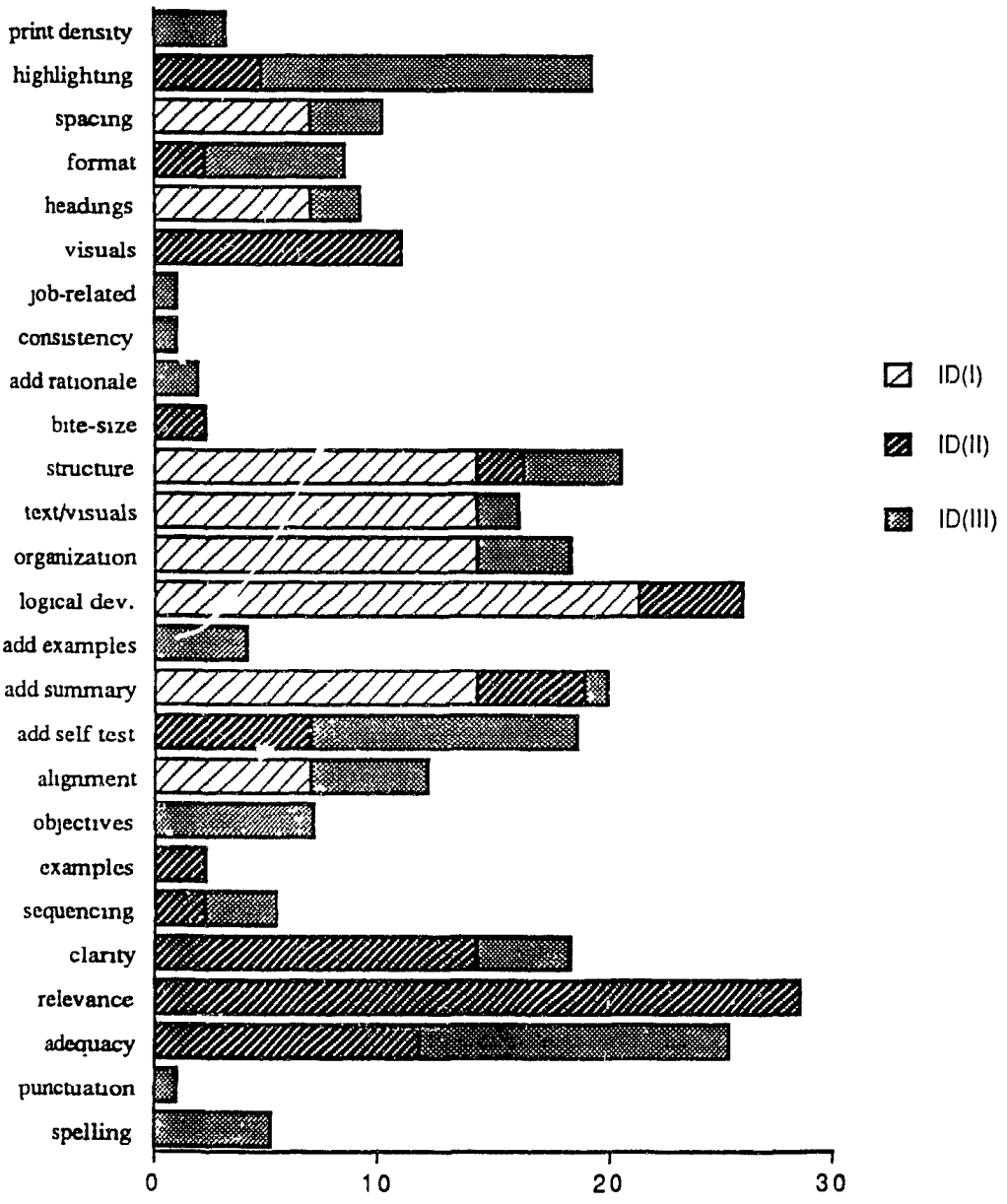

FIGURE 10. **Percentage of type of revision statement by instructional designer**

Knowledge Statements.

Knowledge Statements represented procedural or declarative knowledge of the expert reviewer. More often than not, Knowledge Statements appeared either in conjunction with or in lieu of Problem Identification statements in order to substantiate a detected weakness in the text. Alternatively, they functioned as a Rationale for the proposed revision. Table 15 displays the between groups mean data on Knowledge Statements.

Table 15

Between Groups Mean Data on Knowledge Statements

	Content Experts	Instructional Designers
$\bar{X}$ Frequency	50.3	29.3
$\bar{X}$ Percent (S.D.)	18.5 (10.7)	5.8 (4.7)
$\bar{X}$ % Content	96.4	38.1
$\bar{X}$ % Design	0.0	15.3
$\bar{X}$ % Presentation	0.0	5.6
$\bar{X}$ % Pedagogy	3.5	40.7

As the results indicate, the mean percentage of Knowledge Statements is considerably higher for the Content Experts ($\bar{X}$=18.5 versus $\bar{X}$=5.8). The pattern of distribution among the subcategories observed in Problem Identification and Revision Statements, was also present in Knowledge Statements. While it was expected that the Content Experts would refer to declarative knowledge in order to

substantiate their comments, it was an unexpected finding that the Instructional Designers would also resort to Content and Pedagogical Knowledge for the same reason.

Within group data on this category are presented in Tables 16 and 17.

Table 16

Percentage of Type of Knowledge Statement by Content Expert

	Content Experts		
	I	II	III
Content	100.0	100.0	89.4
Design	0.0	0.0	0.0
Presentation	0.0	0.0	0.0
Pedagogy	0.0	0.0	10.6

Table 17

Percentage of Type of Knowledge Statement by Instructional Designer

	Instructional Designers		
	I	II	III
Content	20.0	58.0	36.3
Design	20.0	8.0	18.1
Presentation	0.0	8.0	9.0
Pedagogy	60.0	26.0	36.3

As is apparent, Knowledge Statements of all the Content Experts were overwhelmingly subject matter related, while those of the Instructional Designers were drawn from other domains, namely content, design, presentation, pedagogy. Of particular interest is the fact that none of the Instructional Designers hesitated to refer to domain knowledge in order to support their point of view.

In summary, the profile of the category based on which goals and criteria were set, change implemented, and scope of the operation monitored, suggested the following: the body of knowledge for the Content Expert was domain related, while for the Instructional Designer, it was mainly the Instructional Development Model. Figure 11 summarizes the proportion of type of knowledge across subjects.

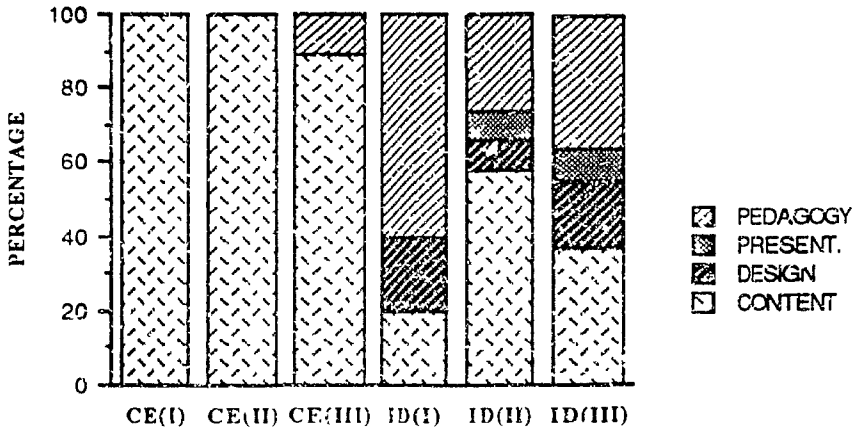

FIGURE 11.

Percentage of Knowledge Statement by Expert

Text Knowledge.

Segments which represented knowledge acquired from the text, were coded as Text Knowledge. Group data on this category are presented in Table 18.

Table 18

Between Groups Mean Data on Text Knowledge

	Content Experts	Instructional Designers
$\bar{X}$ Frequency	5	20.6
$\bar{X}$ Percent (S.D.)	1.2 (1.0)	5.4 (2.9)

Results indicate that the proportion of Text Knowledge was higher for the Instructional Designers than the Content Experts. This finding was expected as Text Knowledge represented knowledge acquired from the stimulus text, and it was the designer group that was not familiar with the subject matter.

Table 19 depicts within group results on this category.

Table 19

Percentage of Text Knowledge by Expert

	(I)	(II)	(III)
Content Expert	2.0	0.0	1.7
Instructional Designer	7.8	6.3	2.1

Despite a higher variability among the Instructional Designers (Min. 2.1, Max. 7.8), the presence of this category in all three subjects' protocols, coupled with points of occurrence of these statements in the protocols suggest that the interaction of the designers with the text also comprised a learning component. In other words, Text Knowledge Statements appeared when the subject attempted to answer the self-test questions by referring to information acquired from the text. Performance on the test, which suggested the amount learned, indirectly provided a gauge for evaluating the material.

Verbatim Statements.

This category represented those parts of the stimulus text which were reproduced verbatim, and it comprised large portions of almost all subjects' protocol. Group data on this category are presented in Table 20.

Table 20

Between Groups Mean Data on Verbatim Statements

	Content Experts	Instructional Designers
$\bar{X}$ Frequency	82.3	80.0
$\bar{X}$ Percent (S.D.)	25.8 (4.1)	22.2 (15.0)

Although mean percentages among the two groups did not differ greatly, the standard deviation of the Instructional Designers was considerably higher than their counterparts.

Within group results are presented in Table 21.

Table 21

Percentage of Verbatim Statement by Expert

	(I)	(II)	(III)
Content Expert	29.0	27.4	21.1
Instructional Designer	38.8	18.6	9.3

Results confirm that there was a closer range in the percentage scores across Content Experts than Instructional Designers (Min. 21.1, Max. 29.0 versus Min. 9.3, Max. 38.8).

An interesting finding was that Verbatim Statements had different functions for each group. This was evident from the sequence of occurrence of these statements throughout the protocols. For the Content Experts, Verbatim statements served as data input. In other words, all comments, including Evaluation, Problem Identification, and Revision Statements were always preceded by a Verbatim Statement. Hence, a section was read, and comments were made in reference to that section. Verbatim Statements for the designers had a secondary function. They represented a strategy of comparison. This implied that the text was not necessarily read in sequence, but rather, sections from

various parts were read subsequent to one another in order to ascertain that the given information was consistent throughout, and sufficient enough to achieve the objectives, and perform well on the self-test.

Indeed, a greater variability in the Verbatim Statements of the Instructional Designers confirms that dependence on input from the text varied among this group members.

Text Talk.

The category of Text Talk represented either a paraphrase, or reference to a specific section of the text. Data on this category are presented in Table 22.

Table 22

Between Groups Mean Data on Text Talk

	Content Experts	Instructional Designers
$\bar{X}$ Frequency	20.0	34.3
$\bar{X}$ Percent (S.D.)	5.3 (2.5)	7.7 (1.9)

The results indicate that Text Talk comprised a higher percentage of the Instructional Designers' protocols. This outcome suggests that the Content Experts, being familiar with the topic, did not perceive it necessary to paraphrase the text for better comprehension.

Within group results for this category is presented in Table 23.

Table 23

Percentage of Text Talk by Expert

	(I)	(II)	(III)
Content Expert	5.9	2.5	7.6
Instructional Designer	8.3	9.3	5.5

The summary of results pertaining to Text Talk, Verbatim Statements, and Text Knowledge are presented in Figure 12.

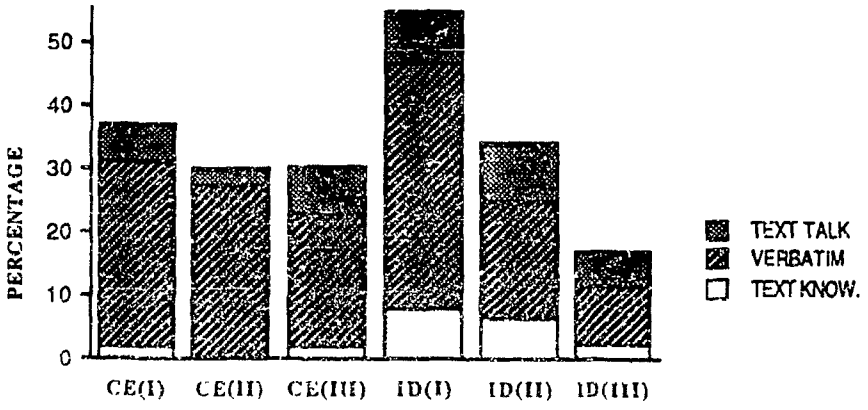

FIGURE 12.

Percentage of Text Talk, Verbatim Statements. and Text Knowledge by expert

Task Related Categories

Task Talk.

Segments referring to action or performance were coded as Task Talk. Specifically, they included: a) the verbalization of the activity that was being carried out; b) a statement requesting a short term goal marked by a 'let's see if...' phrase; c) metacognitive or self-directed question, intended to give direction to the subject's subsequent action (e.g., 'where was I?). Group data on this category are presented in Table 24.

Table 24

Between Groups Mean Data on Task Talk

	Content Experts	Instructional Designers
$\bar{X}$ Frequency	15.6	49.6
$\bar{X}$ Percent (S.D.)	4.2 (1.6)	7.8 (6.1)

Results indicated a larger mean percentage representing Task Talk for the Instructional Designers ($\bar{X}$=7.8>$\bar{X}$=4.2).

Within group results are presented in Table 25.

Table 25

Percentage of Task Talk by Expert

	(I)	(II)	(III)
Content Expert	4.4	2.5	5.8
Instructional Designer	3.7	5.0	14.9

The information conveyed by Task Talk Statements also suggested that the performance of each group was based on a distinctly different frame of reference: one, the domain, and the other, instructional design heuristics. Specifically, tasks verbalized by the Content Experts included a search for definitions or elaborations of concepts. The designers' Task Talk made reference to alignment between parts, consistency in presentation, the presence of adequate number of examples, self-test questions, and summaries.

Strategy Talk.

Performance related statements which referred to a routine course of action, as opposed to an activity which was specific to this task environment, were coded as Strategy Talk. These segments represented some of the activities which the subjects undertook regularly as expert reviewers. Between group data are presented in Table 26.

Table 26

Between Groups Mean Data on Strategy Talk

	Content Experts	Instructional Designers
$\bar{X}$ Frequency	0.3	9.3
$\bar{X}$ Percent (S.D.)	0.1 (0.2)	1.9 (0.2)

Table 27 contains within group results.

Table 27

Percentage of Strategy Talk by Expert

	(I)	(II)	(III)
Content Expert	0.0	0.5	0.0
Instructional Designer	1.8	2.2	1.7

While overall, Strategy Talk did not represent a large proportion of any of the protocols, the number was even smaller for the Content Experts. This, perhaps was due to the fact that the performance of the members of this group was directly tied in with the text rather than governed by a particular set of heuristics. In contrast, the larger number of Strategy Talk comments by the Instructional Designers explicitly referred to evaluation strategies prescribed by the Instructional Design model. This finding suggests that among designers, there is a degree of consistency on using the Instructional Design Model as a basis for formative evaluation. The summary of results on Task Talk and Strategy Talk are presented in Figure 13.

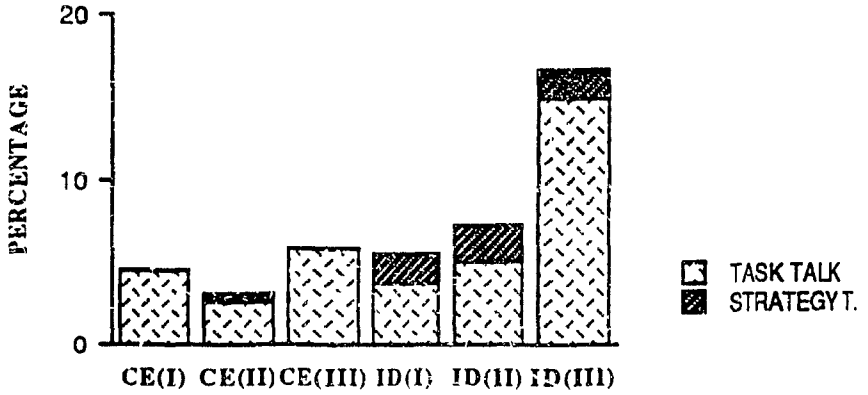

FIGURE 13.

Percentage of Task Talk, and Strategy Talk by expert

Convergence and Divergence on Problem Identification and Revision Statements

In this section, results on convergence of Problem Identification, and Revision Statements across subjects will be presented, and will be followed by a discussion. These data which are presented in Table 28, demonstrates the frequency with which a) the same problem is identified, b) different problems are identified for the same segment, c) similar revision recommendations are made for the same problem, and d) different revision recommendations are made for the same problem.

The numbers in the first column of Table 28 correspond to those segments of the stimulus text (Appendix II) on which two or more subjects have commented. Problem Identification and Revision Statements are depicted by white and black shapes, respectively. Repetition of identical shapes and colours in a row indicate convergence.

Results relating to the Problem Identification of the subjects are presented in Table 29. 'Ratio' represents the number of subjects which made a comment (left hand numbers refer to Content Experts and right hand ones to Instructional Designers).

Table 28
Convergence in Problem Identification and Revision by Expert

Seg. Num.	CE(I)	CE(II)	CE(III)	ID(I)	ID(II)	ID(III)
4-25		□■		○	○	○●
15-20		○●				○■
22		■				△▲
24	●					○●
25	■					△▲
27-31	□			△▲		
33	■				▲	
35	●	○	○			
37	■	○●				○●
38-40	○	○	●		■	
41-44	□■					
47-50	●	△	⬬		○●	○■
50			□			▲
51-53	○	△▲	○●		●	
64-65	●		◕			
72	□			○●	○●	△▲
74	○●	●	○●			○●
75	◕		●			
76-82					●	●
82	▲					□■
93-94			□■		●	●
96	▲	■			⬭	
109	○●		○			
110-113	○●	○●	△●		○■	
114-116	○●				○◔	
123	○	△	○●			○●
124-125	●	○				
128	○	○●	○●		○●	
130		○	○●	□	○●	
136			○		□	△◍
138	■					
145-146	●		●			
147-148				■		△▲
151			○●	○		○●
152				○	○	
154-156	▲		□■			
173			○	●		

Table 29

Between Groups Convergence and Divergence on Problem Identification

	Total	Ratio			
		1/1	1/2	2/1	3/1
Convergence	9	3	1	4	1
Divergence	5	1	1	0	3

Points of convergence between the two groups on Problem Identification were on inadequate content, typographical errors, insufficient or redundant information, and poor spacing and formatting. Divergence was brought forth by the Content Experts' concentration on content matters and the designers' on overall inadequacies.

Within group convergence was overall higher among the Content Experts (8>5) (see Table 28). Interestingly, in both groups, only one problem was referred to by all three subjects. Among the Content Experts, two concurred on problems regarding erroneous and inadequate content, typographical errors, wordiness, and improper use of terms, while two of their counterparts agreed on inconsistencies between the text and objectives, erroneous or unclear content, and poor formatting.

Divergence was slightly lower among the Content Experts (3<4). Disagreement was mainly due to the degree of specificity of the identified problem. For instance where one subject designated poor headings as a problem, the other

highlighted overall inconsistency in structure and presentation.

Convergence and Divergence on Revision Statements

Between group results regarding convergence on Revision Statements are presented in Table 30.

Table 30

Between Groups Convergence and Divergence on Revision

	Total	Ratio				
		1/1	1/2	2/1	3/1	2/1/1
Convergence	10	8	0	1	1	0
Divergence	12	7	3	0	1	1

Aspects on which there was convergence across groups included rearrangement, correction of typographical errors, addition of content, substitution of the format, and adherence to one format. Disagreement arose from tendencies on the part of the Content Experts to focus on local revision suggestions and for the Instructional Designers to make global revision suggestions.

Within group convergence was higher among the Content Experts (6>3), although convergence among all three Designers was more frequent (3>2). Points of agreement for the Content Expert group included correction of typographical errors, addition of content, and substitution of terms. For the Instructional Designers, they included

substitution and addition of content and self-tests.

Within group divergence was also higher among the Content Experts (5>2). When either Problem Identification or Revision was considered as an index for assessing convergence, there was a slight increase in the total number of times where either groups or individuals converged. These data are presented in Table 31.

Table 31

Between and Within Groups Convergence on Problem Identification and Revision

	Total	Subjects		
		4	3	2
Between Groups	13	3	5	5
Within Content Experts	12	-	5	7
Within Group Instructional Designers	7	-	1	6

These results reiterated the point that each expert group focussed on different features of the text. Even within groups, the number of times which group members agreed on a particular instance, was infrequent when it was compared with the frequency data on Problem Identification and Revision (see Tables 1 and 2, Appendix VII).

Revision Statements

Type and Level of Revision.

In order to classify Revision Statements by the type

and level of change, the classification system utilized by Chomsky in group transformations, and which is typically referred to in writing/revision research, (e.g., Faigley & Witte, 1981; Sommers, 1980), was selected. This system categorizes revisions according to the type of change (addition, deletion, substitution, and rearrangement), and the level of change (local and global). Local, also called "surface" change (Faigley & Witte, 1981) is comprised of most copy-editing conventions and other changes which leave meaning intact. They are generally contained within the word or sentence level. Global or "text-base" changes may have either minor or major overall consequences for the text, and may range from adding new examples, to giving the text an entirely new direction. When revisions of subjects were classified with this system, the following results emerged (Table 32).

The results suggest that the Content Experts concentrated on one section of the stimulus text at a time. Their comments, were mostly related to the factual information conveyed by the instructional text. The Designers, on the contrary, made global recommendations. This suggests that their scope of action extended beyond the section they were addressing at any given time. The most common change recommendation appeared to be the addition of various components to the text in order to facilitate learning.

Table 32

Between Groups Percentage of Type of Revision Change

	Content Expert	Instructional Designer
LOCAL CHANGES		
$\bar{X}$ % Addition	45.1	26.9
$\bar{X}$ % Deletion	9.0	1.0
$\bar{X}$ % Substitution	34.8	3.7
$\bar{X}$ % Rearrangement	9.3	4.3
GLOBAL CHANGES		
$\bar{X}$ % Addition	1.3	27.7
$\bar{X}$ % Deletion	0.0	2.7
$\bar{X}$ % Substitution	0.0	11.7
$\bar{X}$ % Rearrangement	0.0	21.5

Within group results are presented in Table 33.

Table 33

Within Groups Percentages of Type of Change

	Content Expert			Instructional Designer		
	I	II	III	I	II	III
LOCAL CHANGES						
Addition	42.8	13.6	79.1	0.0	47.6	33.3
Deletion	14.2	4.5	8.3	0.0	0.0	3.1
Substitution	16.6	81.8	6.2	0.0	7.1	4.1
Rearrangement	26.1	0.0	2.0	0.0	4.7	8.3
GLOBAL CHANGES						
Addition	0.0	0.0	4.1	28.5	21.4	33.3
Deletion	0.0	0.0	0.0	7.1	0.0	1.0
Substitution	0.0	0.0	0.0	14.2	9.5	11.5
Rearrangement	0.0	0.0	0.0	50.0	9.5	5.2

These results confirm that group data represented individual data fairly. The Content Experts nearly always focussed on local changes while the Instructional Designers commented both on the local and global levels.

Revision Statement Components.

In order to assess the degree of specificity of the revision recommendation, these statements were categorized as having a Revision Action, and Revision Result, and a Revision Rationale. Group means representing Revision Statements which ranged from unspecified change to change which was explicit and included a rationale, are presented in Table 34.

Table 34

Between Groups Mean Data on Percentage of Revision Components

	Content Expert	Instructional Designer
Action	7.9	9.4
Action/ Result	11.0	4.7
Action/ Rationale	6.1	34.7
Action/ Result/ Rationale	74.8	50.8

Results suggest that the Revision Statements which were most complete, were made more frequently than those which were not explicit and/or did not contain a rationale. The inadequate elaboration on the remaining Revision Statements could have been due to the nature of the experimental task which required the subjects to think aloud.

Within group results are presented in Table 35.

Table 35

Within Groups Percentage of Revision Statement Components

	Content Expert			Instructional Designer		
	I	II	III	I	II	III
Action	23.8	0.0	0.0	14.2	14.2	0.0
Action/ Result	28.5	4.5	0.0	14.2	0.0	0.0
Action/ Rationale	14.2	0.0	4.1	42.8	42.8	18.7
Action/ Result/ Rationale	33.3	95.4	95.8	28.5	42.8	81.2

As evident, the same trend persisted across subjects. With one exception (Instructional Designer II), the majority of the statements included a Revision Result, or a Revision Rationale. This pattern was particularly prominent among the Content Experts. This outcome could have been brought forth by the tendency to make revision recommendations on the

local level, and for a specific problem. On the contrary, because the Instructional Designers tended to make global recommendations, necessarily their statements were less specific.

Sequence of Coded Categories

The sequence of coded segments charted the processes of formative evaluation. (Three sections from each protocol are presented in Appendix V. This selection represents the beginning, the middle and the final portions of each protocol.)

The observed processes are summarized below:

CONTENT EXPERTS

1. Begin by reading the first section; suspend reading; evaluate the read section.
2. Use the text as a source of information on which they comment.
3. Restrict the range of Evaluation, Problem Identification, and Revision Statements to the preceding Verbatim Statement, thus functioning on a local level.
4. Maintain a consistent pattern throughout. Pattern consists of a Verbatim Statement, followed Evaluation, Problem Identification and/or Revision Statement.
5. Refer to domain knowledge for setting up goals and criteria for formative evaluation.

INSTRUCTIONAL DESIGNERS

1. Begin by reading or skimming the entire text before making any evaluative comment.

2. Use the text as data input as well as for comparing content.

3. Extend Evaluation, Problem Identification, and Revision Statements to the entire text, thus functioning also on a global level.

4. Begin by an overall evaluation of text, and end by summerizing all the recommendations deemed necessary.

5. Refer to design and pedagogical knowledge for setting up goals and criteria.

General Discussion

The results of the current study suggest that there are differences between and on a general level, similarities within group members. Disparities which were observed in both the outcome, that is the product, and the processes of formative evaluation, will be discussed in this section.

Between Group Differences

The outcome of formative evaluation was depicted by four categories: Evaluation, Problem Identification, Revision, and Knowledge Statements. Two findings were derived from the quantitative attributes of the data on these categories. One of the findings was that each expert group formed distinct task representations, and interpreted their role in formative evaluation in a unique way. The

Content Experts functioned as 'specialists', while the Instructional Designers behaved more like 'generalists'.

The most apparent reason for this tendency might be that formative evaluation is a familiar task for an experienced instructional designer whereas it may be novel activity for a content expert. Instructional designers routinely engage in formative evaluation, sometimes independently and sometimes as coordinator and implementor of revision recommendations which are in turn collected from a team of experts (Bowler, 1978; Nevo, 1977; O'Donnell, 1985). This routine, necessarily, familiarizes them with the types of comments which specialists, including subject matter experts, teachers/trainers, and graphic artists render. In fact, numerous checklists published as job-aids for formative evaluators provide guidelines for addressing aspects of materials which extend beyond design expertise. (See Saroyan & Geis, 1988 for a list of 1009 items compiled from 48 such checklists.)

Content experts, on the other hand, are typically called upon as resource persons and are generally expected to comment on domain rather than design or pedagogical issues. Hence, in interpreting their task, they rely heavily on that aspect of their expertise for which they are usually consulted.

It should be pointed out, however, that this 'generalist' role is not necessarily superior to that of a 'specialist'. Results from this study indicate that comments

made by the Instructional Designers on content issues arose only when these subjects found the described concepts unclear or inadequate. Other significant inadequacies, such as erroneous content were often missed. An example of this disparity was observed in the comments made on one segment by Content and Design Experts. Where all three Content Experts found fault with the usage of Fahrenheit instead of Celsius in an example within a "scientific text", and recommended the conversion of the unit, the instructional designer found the example very 'crisp' and effective. Such evidence lends support to the assertion that instructional designers do not necessarily identify the same inadequacies in the text as content experts. They are particularly ill-equipped to identify erroneous content, and are therefore, unable to make appropriate revisions.

This finding has significant practical implications. A limited task representation on the part of the Content Expert may inhibit this type of expert from fulfilling requirements of a successful formative evaluator. Conversely, instructional designers as 'generalists', may and will attempt to present facts and concepts in a systematic and consistent fashion, and may be able to make the text more appealing to the learner by rearranging the format, headings, and visuals. They may further be able to identify complications in the narrative. Yet they fall short in specifying content errors and, moreover, are unable to fix them.

The stimulus text used in this study is a good example of a text which was initially developed by instructional designers with some initial input from subject matter experts. This product clearly demonstrates that when facts are manipulated and presented as a discourse by instructional designers who do not necessarily possess a knowledge of the subject matter, they may get distorted and become unacceptable to the content expert revisers in a second round review. In forming a team of formative evaluators and projecting a budget for this process, the limitations of using only one kind of expert should be considered. Moreover, since the results of this study indicate that members of either group tend to interpret their task in the same way and on a general level comment on similar issues, it may be more beneficial to have, for instance, one of each group instead of two members of one group in a formative evaluation team.

Other evidence on group differences was observed in the qualitative aspect of comments generated by the members of each group. Disparity was witnessed in Evaluation Statements, Problem Identification, and in the type of Revision Statements. For instance, almost all of the Revision Statements of the Content Experts were local and on the word or sentence level, while those of the Instructional Designers were also for the overall improvement of the text.

The particular finding regarding Revision Statements is interesting in that, in a way, it resembles other studies

which have compared revisions of expert and novice writers (Bartlett & Scribner, 1981; Flower & Hayes, 1977; Markman, 1977; NAEP, 1977; Sommers, 1980; Stallard, 1974). They have concluded that when making changes, skilled revisers first revise globally and then locally, while the unskilled, novice revisers attend to the word level.

Results of the current study indicate that the revision of the Content Experts is similar to novices, while the Instructional Designers perform in a manner which closely resembles the expert reviser. As suggested earlier, it is possible that this difference is closely tied to the degree of experience in formative evaluation, i.e., review and revision.

The data also pointed to the distinction between the two groups in the strategies they adopted to perform the formative evaluation task. Three sets of strategies were identified as being representative of one of the two groups:

a) content versus heuristic driven strategy;

b) sequential versus comparative strategy;

c) 'expert' versus 'naive learner' strategy.

A: Content versus Heuristic Driven Strategy

The results of several sets of data suggested that each of the expert groups depended upon a distinctly different heuristic in guiding them through the task. While the Content Experts based their evaluation on heuristics dictated by the domain, the Instructional Designers used the

Instructional Development (ID) model and the heuristics recommended by it to evaluate the text. Evaluation, Problem Identification, and Revision Statements suggested that Instructional Designers carry out their evaluation by assessing the degree of adherence of the text to the prescriptions of the ID model. One major criterion in this evaluation appeared to be the alignment of various parts of the text. Another was the validity of both the course and lesson objectives in terms of accomplishing the set goals.

These standards concur with recommendations which have appeared in the instructional design literature regarding the development of sound instructional material. For instance, Burkholder, (1981-82), Choi et al., (1979), Merrill et al., (1979) have cited 'consistency' and 'adequacy' as two of the most important attributes. Other recommended criteria include: a) adhering to presentation principles in order to enhance learning such as effective highlighting, b) maintaining consistency in typeface, c) avoiding right justification, and d) placing visuals within a close proximity to the text (Felker, Redish, & Peterson, 1985; Hartley, 1978, 1981; Smillie, 1985; Wright, 1985).

This finding is significant in that it suggests that Instructional Designers function more on the basis of a system rather than intuition. Nonetheless, it also points to a disadvantage. Basing evaluation on consistency and adequacy presumes that the content is accurate and relevant. While it is interesting that despite the lack of domain

knowledge, all three Instructional Designers commented on this issue, it emphasizes shortcomings when it comes to detecting erroneous or irrelevant content. Data on Task Talk and Strategy Talk also confirmed that the ID model had a direction giving function for the Designers. The short term goals coded as Task Talk confirmed that Content Experts searched for definitions or elaborations of concepts. The designers looked for such things as alignment between parts, consistency in presentation, and the presence of adequate number of examples, self-test questions, summaries and overviews.

The near absence of Strategy Talk from the protocols of the Content Experts, and the verbalized strategies by the Instructional Designers clearly indicated that the latter group functioned on the basis of a system rather than reacting spontaneously to the stimulus text.

Additionally, a sequence or chain of Revision Statements in the Instructional Designers' protocols confirms that this group carried out their task on the basis of a plan, and not intuition.

B: Sequential versus Comparative Strategy

Closely tied with the strategies discussed above is the way in which the stimulus text was used by each group. Data on Verbatim Statements suggest that the Content Experts interacted with the text sequentially, thus using it as the source of data for formative evaluation. The Instructional

Designers, on the other hand, used the text comparatively. This implied that the text was not necessarily read in sequence, but rather, sections from various parts were read subsequent to one another in order to ascertain that the given information was consistent throughout, and sufficient enough to achieve the objectives.

C: Expert versus Naive Learner Strategy

As discussed in the previous section, the Content Experts assumed a 'specialist' role as compared with the 'generalist' role of the Instructional Designers. In addition, the latter group members assumed a secondary role as a naive learner, and attempted to study the content as a potential user would. In doing this, they strived to identify and anticipate the difficulties which a naive learner was bound to encounter.

Several sets of the data lend support to this view. For instance, data on Text Knowledge which represents information gained from the text suggest that there is a higher percentage of Text Knowledge Statements for Instructional Designers than for Content Experts. Furthermore, as a response to the test questions, Text Knowledge Statements appear only in the protocols of the Instructional Designers. The Content Experts, on the other hand, evaluated questions at their face value rather than using them as an additional scale for assessing the effectiveness of the text. Finally, a number of the

Evaluation Statements of the Instructional Designers represent an evaluation of personal performance on the test. This is also perceived as an indication that a successful performance on the test translates into a favourable, although indirect, evaluation for instructional designers.

A Process Model of Formative Evaluation

The processes observed in the formative evaluation of Content Experts and Instructional Designers are quite similar to the revision processes of expert writers which was described in the Literature Review section. The Process Model of Revision, conceptualized by Hayes et al. (1987), (see page 24), can be slightly modified in order to describe the behaviour of the experts of the present study. Primarily, where **Read** is branched out into **Comprehend**, **Evaluate**, and **Define Problems**, the processes could be modified to **Comprehend , Compare and / or Evaluate** to accommodate the strategy which typically the Instructional Designers applied. Secondly, although Evaluation (as defined in this study, and more specifically Problem Identification are processes which necessarily take place before a modification is made, their verbalization may occur after the revision itself. For instance, after reading a text which contains a typographical error, the subject may say: *I'm going to change 'and' to 'an'/ the 'd' is extra/.* In such a case, the sequence of the processes presented in the model does not adequately represent the observed

processes. A more suitable sequence for the Processes is suggested in Figure 14.

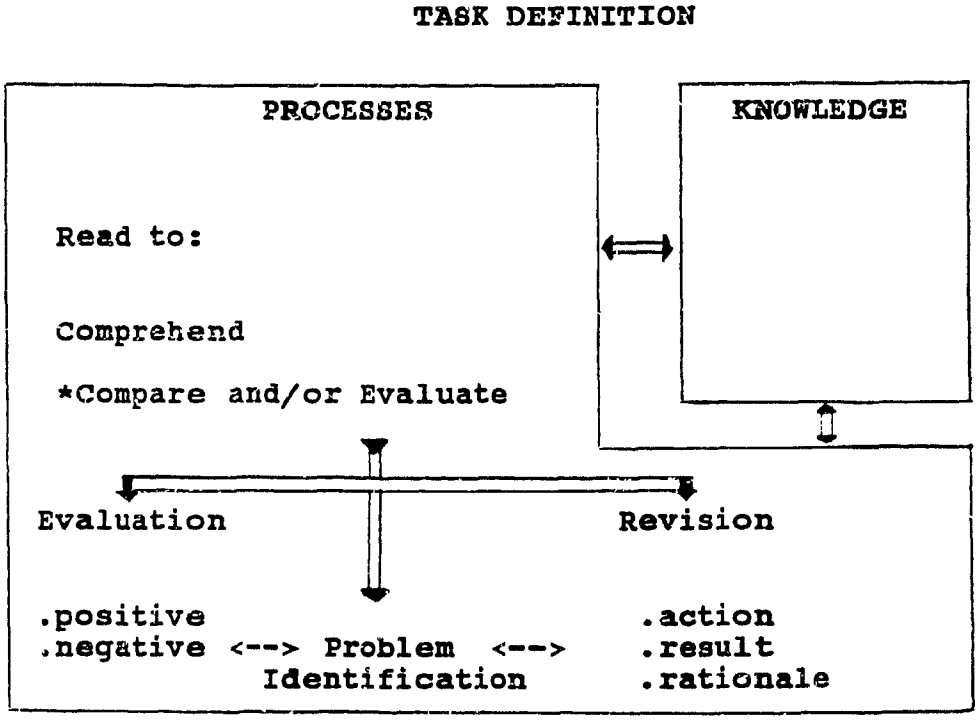

FIGURE 14. Processes of formative evaluation

This sequence suggests that the verbalization of any of the three types of comments could be made immediately after interacting with the stimulus text. These comments are not bound in their sequence, nor are they mutually inclusive. In other words, Evaluation, Problem Identification, and Revision Statements may be made any time, and may appear alone or in conjunction with either or both of the other two comments. The Knowledge component of the model adequately

represents the influence of the particular domain knowledge in setting goals and the criteria for an acceptable formative evaluation.

Within Group Similarities and Differences

Resemblance among group members was observable in the strategies employed rather than the outcome of formative evaluation. In the latter instance, resemblance was confined to a superficial level. For instance, as a rule, the comments of the Content Experts were mostly Content related while the Instructional Designers commented on various aspects of the text.

On a deeper level, the disparity among group members became more apparent. As the data indicated, there was generally a higher range of variability within the Instructional Designer than the Content Expert group. The Evaluation data suggest that the degree of dependency on various aspects of ID model, which by itself, was one common factor among the designers, varied among the three subjects, with about equal amount of concurrence between any two of the three experts. This was apparent from the different weighting on Problem Identification and Revision Statements. The variety in the specific comments in each of these two categories suggests strong idiosyncratic tendencies in addressing particular features of the text.

Convergence data on Problem Identification and Revision suggest that nearly always the same erroneous or inadequate content was detected by the Content Experts. Similarly,

inconsistencies between the objectives and text or test items were pointed out by the Instructional Designers. However, other types of problems were less frequently detected unanimously.

In sum, the results have helped define the processes of formative evaluation and have suggested that Content and Instructional Experts use different strategies to carry out this task. Furthermore, on a superficial level, there are similarities in the types of comments within expert groups. However, comments become less similar when analyzed on a deeper level.

The conclusions drawn from the results and the implications of the findings will be presented in the next chapter.

CHAPTER V

CONCLUSION AND RESEARCH AND PRACTICAL IMPLICATIONS

Several conclusions are drawn from this study regarding the process as well as the product of formative evaluation. On the basis of the qualitative data, the process of formative evaluation, as carried out by Content and Instructional Design Experts, can be described using the following terminology: **Evaluation, Problem Identification, Revision, Knowledge Statements, Text Talk, Verbatim Statements, Text Knowledge, Task Talk, and Strategy Talk.** Future data from studies which include other types of experts can ascertain whether or not this terminology is indeed the most comprehensive and parsimonious definition of the processes of formative evaluation.

Previous research has purported that much of the outcome of formative evaluation is dependent upon the skill and intuition of the expert instructional designer. The results of this study indicate that this statement is partly true. Specifically, it is argued that the Instructional Designers' performance is undeniably linked to, and driven by the heuristics of the instructional design model. Furthermore, the performance of the Content Experts appears to be governed by domain knowledge. Both sources of knowledge, in particular the design heuristics, provide a point of reference to evaluate the stimulus text, and to

identify existing weaknesses. This results in greater consistency within group members in evaluation and problem identification. However, in making subsequent revision recommendations, the instructional design model does not provide revision heuristics. Thus, when evaluations and problem identifications are translated into actual revision, within group differences are likely to be more salient among instructional designers than among content experts. The strong reliance on respective areas of expertise also seems to influence the selection of similar strategies to carry out formative evaluation by members of each group.

It has been argued that there are further qualitative differences between Content Experts and Instructional Designers. Differences are more apparent in the Evaluation, Problem Identification, Revision, and Knowledge Statements. More specifically, it is contended that:

1- Content Experts generally comment on content-related issues while Instructional Designers comment on other aspects, including the design, content, presentation, and pedagogical value of the text. This "specialist" versus "generalist" role confirms previous assumptions which have been made in the literature.

2- Content Experts may, despite their pedagogical experience and familiarity with the target audience, refrain from commenting on the pedagogical aspect of the text. Conversely, it appears that Instructional Designers place emphasis on pedagogical issues and the needs of the potential

user. In fact, as observed in this study, instructional designers may assume the role of a 'naive learner', and may use this technique to evaluate the effectiveness of the text.

3- Content Experts appear to direct their revision recommendations to sections which they are currently evaluating; in other words, they revise locally. Instructional Designers, in contrast, do not seem to be bound to a specific section. They tend to render recommendations for revision which are global in nature and have overall implications.

4- text features which seem to consistently draw the attention of Content Experts are: typography, redundancy, inadequacy, ambiguity of concepts and terms, and the appropriateness of examples. Instructional Designers, on the other hand, tend to notice weaknesses in the design, formatting, organization, consistency, instructional quality, and the appropriateness of the objectives and exercises within the text.

5- the degree of convergence on any Problem Identification and Revision comments, between two or more experts more (Content and Instructional Design) is less than 25% of the time.

6- as a group, Content Experts seem to converge more with their colleagues than the Instructional Designers do with theirs. While the former group is more likely to concur on comments regarding the inaccuracy or inadequacy of content, the Instructional Designers converge on issues such

as inconsistencies between various parts of the text, and poor formatting.

Given the importance of formative evaluation and its potential effect on the improvement of instructional text, the findings of this study have broad implications for future research as well as practice.

The vast majority of the empirical studies on formative evaluation have been product -or outcome-oriented. In order to understand a phenomenon such as formative evaluation completely, it becomes evident from results of this study that in addition to the product, the processes of this activity also need to be characterized. The study of formative evaluation in a problem-solving context has also provided a model for future, theory based research in this area. The innovative application of the think-aloud method of data collection, and the development of a comprehensive, reliable coding scheme, have possibly provided the necessary tools to pursue this line of research. Future replications of this study could add further credibility to this methodology. Other investigations may introduce variations in terms of types of experts or the medium of instruction used as the stimulus text.

In the present study, an initial attempt was made to conceptualize a process model of formative evaluation. However, this model could be further tested and advanced with data from other experts sources as they perform formative evaluation.

The detailed qualitative analysis of the data and the conclusions drawn herein, have potential implications for the practitioner who conducts formative evaluation. It has been concluded that content experts and instructional designers comment on different aspects of the text. Furthermore, the exact nature of these comments have been explicated. This information allows the practitioner to gain insight regarding the scope and limitations of comments generated by content and instructional designers in a formative evaluation project. It has further been contended that the inclusion of both types of experts as a team is more likely to yield results superior to the separate involvement of each type of expert in the same project. This knowledge can reduce duplication in effort, can increase efficiency, and can yield an optimal product which is mutually acceptable.

This study also has concluded that on a more microscopic level, the comments within groups will not be necessarily similar. This finding may raise this question for the practitioner: What is a viable number of experts from each group in order to elicit the most comprehensive set of comments regarding problem identification and revision? It is important to remember that while comments vary greatly when examined on a very detailed level, either set of experts, consistently comment on specific issues. In practical terms, several instructional designers or content experts will most likely concur on gross inadequacies in the text determined by their frame of reference. Thus, in case of budget and/or

time restrictions, it would seem preferable to select evaluators representing different fields of expertise rather than several from one group.

Finally, the notion that content experts with considerable teaching expertise will not necessarily comment on the pedagogical value of the text is a sobering finding which has far reaching educational and practical implications. In past and present practice, many textbooks have been written and reviewed by subject matter experts alone. It has been argued that in order to ascertain the pedagogical value of instructional materials, it is eminently important to consult with other sources of expertise which are more sensitive towards the needs of the students, before the product is published.

In sum, this study has described the processes of formative evaluation as observed in the performance of a review task by content and instructional design experts. Furthermore, it has emphasized the qualitative differences which exist between the two types of experts, and has identified and described levels on which experts are likely to converge and diverge in their comments.

REFERENCES

Abedor, A. J. (1971). Development and validation of a model explicating the formative evaluation process for multi-media self-instructional learning systems. Dissertation Abstracts International, 32,9-10. (University Microfilms No. 72-8621)

Alkin, M. C., & Fitz-Gibbon, C. T. (1975). Methods and theories of evaluation programs. Journal of Research and Development in Education, 8(3), 2-15.

Ally, M. (1985). A team approach to computer courseware design. Educational Technology, 25(6), 29-30.

Anderson, J. (1976). Language, memory, and thought. Hillsdale, NJ: Erlbaum.

Anderson, J. (1982). Acquisition of cognitive skill. Psychological Review, 89, 4, 369-406.

Andrews, D. H., & Goodson, L. A. (1980). A comparative analysis of models of instructional design. Journal of Instructional Development, 3(4), 2-16.

Baghdadi, A. (1980). A comparison between two formative evaluation methods. (Doctoral dissertation, Indiana State University). Dissertation Abstracts International, 41, 8, 3387-A.

Baker, E. L. (1970). Generalizability of rules for an empirical revision. Audiovisual Communication Review, 18, 3, 300-305.

Bamberg, B. (1978). Composition instruction does make a difference. Research in the Teaching of English, 12, 47-59.

Bartlett, E. J., & Scribner, S. (1981). Text and content: An investigation of referential organization in children's written narratives. In C. H. Frederiksen & J. F. Dominic (Eds.), Writing: Process, development and communication (pp. 153-167). Hillsdale, NJ: Erlbaum.

Beach, R. (1976). Self evaluation strategies of extensive revisers and non revisers. College Composition and Communication, 27, 1, 160-162.

Beach, R. (1979). The effects of between-draft teacher evaluation versus student self-evaluation on high school student's revising of rough drafts. Research in the Teaching of English, 13, 111-119.

Beach, R., & Eaton, S. (1984). Factors influencing self-assessing and revising by college freshmen. In R. Beach & L. Bridwell (Eds.), New directions in composition research(149-170). New York: Guilford.

Bell, R. C., & Sullivan, J. L. F. (1981). Student preferences in typography. Programmed Learning and Educational Technology, 18, 2, 57-61.

Black, J., Galambos, J., & Reiser, B. (1984). Coordinating discovery and verification research. In D. Kieras & M. Just (Eds.), New methods in the study of immediate processes in comprehension (237-297). Hillsdale, NJ: Erlbaum.

Bowler, M. (1978). The making of a textbook. Learning, 6(March), 38-42.

Bracewell, R. J. (1980). Writing as a cognitive activity. Visible Language, 14, 4, 400-422.

Bracewell, R. J. (1983). Investigating the control of writing skills. In P. Mosenthal, P. L. Tamor & S. Walmsley (Eds.), Research on writing: Principles and methods (177-204). New York: Longman.

Branson, R. K. (1973). Analysis and assessment of the state of the art in instructional technology. Fort Monroe, VA: Army Training and Doctrine Command. (NTIS Document Reproduction Service No. ADA010 394: Eric Document Reproduction Service No. ED 088-436.

Bridwell, L. S. (1980). Revising strategies in twelfth grade students' transactional writing. Research in the Teaching of English, 14, 197-222.

Briggs, L. (Ed.) (1977). Instructional design principles and applications. Englewood Cliffs, NJ: Educational Technology.

Breuleux, A. (1987a, April). Expert writer's planning. Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.

Breuleux, A. (1987b, July). Discourse and the investigation of cognitive skills in complex tasks. Paper presented at the meeting of the Cognitive Science Society, Seattle, WA.

Burkholder, B. L. (1981-82). The effectiveness of using the instructional strategy diagnostic profile to prescribe improvements in self-instructional materials teaching abstract concepts. Journal of Instructional Development, 5(2), 2-9.

Burt, C. W. (1986, April). An experimental study of group size and participant's role in developmental testing. Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.

Burt, C. W. (1989). Identification of critical variables in developmental testing and the experimental examination of the number and roles of participants in testing sessions. Undefended doctoral dissertation, McGill University, Montreal.

Calkins, L. M. (1980). Notes and comments: Children's rewriting strategies. Research in the Teaching of English, 14, 4, 331-341.

Cambre, M. (1981). Historical overview of formative evaluation of instructional media product. Educational Communications and Technology Journal, 29(1), 1-25.

Carroll, M. J. (1988, April). Differences in quantitative and qualitative data from four developmental testing conditions. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. Cognitive Psychology, 4, 55-81.

Chiesi, H. L., Spilich, G. J., & Voss, J. F. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. Journal of Verbal Learning and Verbal Behavior, 18, 257-273.

Choi, S. Y., Merrill, M. D., Callahan, E., Hawkins, R. T., & Norton, R. F. (1979). The relationship of test performance to ISDP rating in organic chemistry texts. Journal of Instructional Development, 3(1), 16-25.

Cooper, R. J., & Flavell, J. H. (1975). Cognitive correlates of children's role-taking behavior. Mimeographed copy: University of Minnesota. (Reference appears in Glusksburg, Kraus, & Higgins, Review of Child Development Research, 4, 305-345)

Cooper, M., & Holtzman, M. (1985). Counterstatement reply. College Composition and Communication, 36, 97-100.

Cowan, J. (1980). Is systematic curriculum design always feasible? Programmed Learning and Educational Technology, 17, 2, 115-117.

Cronbach, L. J. (1963). Course improvement through evaluation. Teachers College Record, 64, 672-683.

Crowhurst, M. (1978). Audience and mode of discourse effects on syntactic complexity at two grade levels (Doctoral dissertation, University of Minnesota, 1977). Dissertation Abstracts International, 38, 12, 7300A.

DAE Project (1978). Microbiology: A submodule to sterilization and disinfection. Seattle, WA: Office of Research in Medical Education, School of Medicine, University of Washington.

de Groot, A. D. (1966). Perception and memory versus thought: Some old ideas and recent findings. In B. Kleinmuntz (Ed.), Problem solving. New York: Wiley.

Deisler, E. P., & Mc Neil, J. D. (1960). Progress report, video disc instructional program. (ERIC Document Reproduction Service No. 1ED 002 558)

Dick, W. (1968). A methodology for the formative evaluation of instructional materials. Journal of Educational Measurement, 5, 2, 99-102.

Dick, W. (1977). Formative evaluation. In L. J. Briggs (Ed.), Instructional design: Principles and applications (pp. 311-333). Englewood Cliffs, NJ: Educational Technology.

Dick, W. (1980). Formative evaluation in instructional development. Journal of Instructional Development, 3(3), 3-6.

Dick, W., & Carey, L. (1985). The systematic design of instruction (2nd ed.). Glenview, IL: Scott, Foresman.

Dillinger, M. (1987). Segmentation and clause analysis (Tech. Rep. No. 3). Montreal, Canada: McGill University, Laboratory for Applied Cognitive Science.

Dodd, B. T., Lehunte, R. J. G., & Sheppard, C. (1975). Decision making in instructional design. In J. P. Baggaley (Ed.), Aspects of educational technology, VIII. London: Kogan Page.

Duchastel, P. C. (1983a). Text illustrations. Performance and Instruction Journal, 22(4), 3-5.

Duchastel, P. C. (1983b). Toward the ideal study guide: An exploration of the functions and components of study guides. British Journal of Educational Technology, 14, 3, 217-231.

Duchastel, P. C., & Whitehead, D. (1980). Exploring student reactions to inserted questions in text. Programmed Learning and Educational Technology, 17, 1, 41-47.

Dupont, D., & Stolovitch, H. (1983). The effects of a systematic revision model on revisers in terms of student outcomes. National Society for Performance and Instruction Journal, 22(2), 33-37.

Egan, D. E., & Schwartz, B. J. (1979). Chunking in recall of symbolic drawings. Memory and Cognition, 7(2), 149-158.

Ellis, J. A., & Wulfeck, W. H., II. (1978). The instructional quality inventory: Vol. 4. Job performance aid (Special Report No. 79-5). San Diego, CA: Navy Personnel Research and Development Center.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. Psychological Review, 87, 3, 215-251.

Ericsson, K. A., & Simon, H. A. (1984). Protocol analysis. Cambridge, MA: MIT.

Faigley, L. (1984). Measuring the effects of revision on text structure. In R. Beach & L. Bridwell (Eds.), New directions in composition research (pp. 95-108). New York: Guilford.

Faigley, L., Cherry, R. D., Jolliffe, D. A., & Skinner, A. (1985). Assessing writer's knowledge and processes of composing. Norwood, NJ: Ablex.

Faigley, L., & Witte, S. (1981). Analyzing revision. College Composition and Communication, 32, 4, 400-414.

Faust, S. M. (1980). Instructional developer as content specialist: Three case studies utilizing the instructional development operations research model. Educational Technology, 20(9), 5-12.

Felker, D., Redish, J., & Peterson, J. (1985). Training authors of informative documents. In T. Duffy & R. Waller (Eds.), Designing usable text (pp. 43-61). Orlando, FL: Academic.

Flavell, J. H. (1963). The developmental psychology of Jean Piaget. NJ: Van Nostrand.

Flavell, J. H. (1977). Cognitive development. Englewood Cliffs, NJ: Prentice-Hall.

Fleming, M., & Levie, W. H. (1978). Instructional message design. Englewood Cliffs, NJ: Educational Technology.

Flower, L. (1979). Writer-based prose: A cognitive basis for problems in writing. College English, 41, 1, 19-38.

Flower, L., Carey, L., & Hayes, J. R. (1985). Diagnosis and revision. The experts options (Tech. Rep. No. 27). Pittsburgh, PA: Carnegie-Mellon University, Communications Design Center.

Flower, L., & Hayes, J. R. (1977). Problem solving strategies and the writing process. College English, 39, 4, 449-461.

Flower, L., & Hayes, J. R. (1981a). A cognitive process theory of writing. College Composition and Communication, 32, 4, 365-387.

Flower, L., & Hayes, J. R. (1981b). Plans that guide the composing process. In C. H. Frederiksen & J. F. Dominic (Eds.), Writing: The nature, development and teaching of written communication: Vol. 2 (pp. 39-58). Hillsdale, NJ: Erlbaum.

Flower, L., Hayes, J., Carey, L., Schriver, L., & Stratman, J. (1986). Detection, diagnosis and the strategies of revision. College Composition and Communication, 37, 1, 16-55.

Flower, L., Hayes, J. R., & Swarts, H. (1980) Revising functional documents: The Scenario Principle. (Tech. Rep. No. 10). Pittsburgh: Carnegie-Mellon University, Document Design Project.

Foshay, W. R. (1984). QA and QC: A training vendor's vision of the formative/summative evaluation distinction. Performance and Instruction Journal, 23(10), 15-17.

Frase, L. E. De Gracie, J. S., & Poston, W. K., Jr. (1974). Product validation: Pilot test or panel review? Educational Technology, XIV(8), 32-35.

Frase, L. T. (1981). Writing, text and the reader. In C. H. Frederiksen & J. F. Dominic (Eds.), Writing: Process, development and communication (pp. 209-221). Hillsdale, NJ: Erlbaum.

Friesen, P. A. (1973). Designing instruction. Santa Monica, CA: Miller.

Gagne, R. M., & Briggs, L. J. (1974). Principles of instructional design (1st ed.). New York: Holt, Rinehart, & Winston.

Gagne, R. M., & Briggs, L. J. (1979). Principles of instructional design (2nd ed.). New York: Holt, Rinehart & Winston.

Geis, G. L. (1986, April). Student input on materials development. Paper presented at the annual meeting of the American Educational Research Association, San Fransisco, CA.

Geis, G. L. (1987). Formative evaluation: Developmental testing and expert review. Performance and Instruction, 26(4), 1-8.

Geis, G. L. (1988, April). Profiles of the actors in formative evaluation. Paper presented at the annual conference of the American Educational Research Association, New Orleans, LA.

Geis, G. L., Burt, C. W., & Weston, C. B. (1984,April). Instructional development: Developmental testing. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Glaser, R. (1985). Thoughts on expertise (Tech. Rep. No. 8). Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center, University of Pittsburgh, PA.

Golas, K. C. (1982). The effectiveness and cost of alternate models of formative evaluation for printed instructional materials (Doctoral dissertation, Florida State University). Dissertation Abstracts International, 43, 2873A.

Golas, K. C. (1983). Formative evaluation effectiveness and cost. Performance and Instruction, 22(5), 17-19.

Gooler, D. D. (1980). Formative evaluation strategies for major instructional development projects. Journal of Instructional Development, 3(3), 7-11.

Greeno, J., & Simon, H. (1984). Problem solving and reasoning (Tech. Rep. No. APS-14). Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center.

Gropper, G. L. (1975). Diagnosis and revision in the development of instructional materials. Englewood Cliffs, NJ: Educational Technology.

Gropper, G. L., Lumsdaine, A. A., & Shipman, V. (1961). Studies in televised instruction: Improvement of televised instruction based on student responses to achievement tests. ERIC ED 033-647.

Halliday, M. (1967a). Notes on transitivity and theme in English: Part I. Journal of Linguistics,3, 37-81.

Halliday, M. (1967b). Notes on transitivity and theme in English: Part II. Journal of Linguistics,3, 177-274.

Halliday, M. (1968). Notes on transitivity and theme in English: Part III. Journal of Linguistics,4, 179-215.

Hartley, J. (1978). Designing instructional text. London: Kogan Page.

Hartley, J. (1981). Eighty ways of improving instructional text. IEEE Transactions on Professional Communications, PC 24, 1, 17-27.

Hartley, J., & Trueman, M. (1981). The effects of changes in layout and changes in wording on preferences for instructional text. Visible Language, 15, 1, 13-31.

Hartley, J., Trueman, M., & Burnhill, P. (1980). Some observations on producing and measuring readable writing. Programmed Learning and Educational Technology, 17(3), 164-174.

Hayes, J. R., & Flower, L. (1980). Writing as problem solving. Visible Language, XIV, 4, 388-399.

Hayes, J. R., & Flower, L. (1983). Uncovering cognitive processes in writing: An introduction to protocol analysis. In P. Mosenthal, L. Tamor, & S. Walmsley (Eds.), Research on writing: Principles and methods (pp. 207-220). New York: Longman.

Hayes, J. R., Flower, L., Schriver, K., Stratman, J., & Carey, L. (1987). Cognitive processes in revision. In S. Rosenberg (Ed.), Advances in applied psycholinguistics: Vol. 2 (pp. 176-240). Cambridge, England: Press Syndicate of the University of Cambridge.

Hayes, J. R., & Simon, H. A. (1974). Understanding written problem instructions. In L. W. Gregg (Ed.), Knowledge and cognition. Potomac, MD: Erlbaum.

Henderson, E. S., & Nathenson, M. B. (1976a). Developmental testing: An empirical approach to course improvement. Programmed Learning and Educational Technology, 13(4), 31-42.

Henderson, E. S., & Nathenson, M. B. (1976b). Developmental testing: Collecting feedback and transforming it into revision. National Society for Performance and Instruction Journal, 16(3), 6-10.

Henderson, E. S., & Nathenson, M. B. (1977). Case study in the implementation of innovations: A new model for developmental testing. In P. Hills & J. Gilbert (Eds.), Aspects of educational technology(114-120). London: Kogan Page.

Hull, G. (1983). The editing process in writing: A performance study of experts and novices (Doctoral Dissertation, University of Pittsburgh, 1983). Dissertation Abstracts International, 45, 1-3, 829A.

Johnson, S. (1988). Cognitive analysis of expert and novice troubleshooting performance. Performance and Instruction, 1(3), 38-54.

Kandaswamy, S. (1980). Evaluation of instructional materials. A synthesis of models and methods. Educational Technology, 20(6), 19-26.

Kandaswamy, S., Stolovitch, H. D., & Thiagarajan, S. (1976). Learner verification and revision: An experimental comparison of two methods. Audio Visual Communication Review, 24, 3, 316-338.

Kern R. P., Sticht, T. G., Welty, D., & Hauke. R. N. (1976). Guidebook for the development of army training literature. Washington, DC: US Army Institute for the Behavioral and Social Sciences and Human Resources Research Organization

Klare, G. R. (1963). The measurement of readability. Ames, IA: Iowa State University.

Klare, G. R. (1976). Judging readability. Instructional Science, 5, 55-61.

Kniffin, J. D. (1978). The practical application of readability/comprehensibility research to the production of more usable military materials. Hunt Valley, MD: Westinghouse Electric Corporation.

Komoski, K. P. (1971). Statement to the U.S. House of Representatives Committee on Education and Labor. In Hearing: To establish a National Institute of Education, 92nd Congress, First Session. Washington, DC: U.S. Government Printing Office.

Komoski, P. K. (1983). Formative evaluation. Performance and Instruction Journal, 22(5), 3-4.

Komoski, P. K., & Woodward, A. (1985). The continuing need for Learner Verification Revision of textual material. In D. H. Jonassen (Ed.), The technology of text: Vol. 2 (pp. 396-417). Englewood Cliffs, NJ: Educational Technology.

Krippendorff, K. (1980). Content analysis. Beverley Hills, CA: Sage.

Kuipers, B., & Kassirer, J. P. (1984). Causal reasoning on medicine: Analysis of a protocol. Cognitive Science, 8, 363-385.

Larkin, J. (1983). The role of problem representation in physics. In D. Genter & A. Stevens (Eds.), Mental models (pp. 75-98). Hillsdale, NJ: Erlbaum.

Larkin, J. H., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Models of competence in solving physics problems. Cognitive Science, 4, 317-345.

Logan, R. S. (1982). Instructional systems development. New York: Academic.

Lowenthal, D. (1980). Mixing levels of revision. Visible Language, 14, 383-387.

Lumsdaine, A. A. (1966). Recommendations for reporting the effectiveness of programmed instruction materials. Washington, D. C.: Division of Audio Visual Instruciton Services.

Macdonald-Ross, M. (1978). Language in texts. Review of Research in Education, 6, 229-275.

Markle, S. M. (1967). Empirical testing of programs. In Programmed Instruction: Part II (104-138). The 66th annual yearbook of the National Society of the Study of Educational Research. Chicago, IL.

Markle, S. M. (1976). Evaluating instructional programs: How much is enough? National Society for Performance and Instruction Journal, 15(2), 1-5.

Markman, E. (1977). Realizing you don't understand: A preliminary investigation. Child Development, 48, 986-992.

Martelli, M. (1979). A study of a theory-based model of formative evaluation (Doctoral dissertation, Florida State University). Dissertation Abstracts International, 40, 6, 3089-A.

Matsuhashi, A. (1981). Pausing and planning: The tempo of written discourse production. Research in Teaching of English, 15, 113-134.

Mc Entie, R., & Rivers, L. C. (1971). Introduction to the psychology of leadership. An analysis and evaluation of instructional methodology for a multi-media course in leadership. Psychology and Management, Phase III (Evaluation Rep. No. TR-6-15). Annapolis, MD: Naval Academy. (ERIC Document Reference No. ED 07 349)

Merrill, M. D., & Boutwell, R. C. (1973). Instructional development methodology and research. Review of Research in Education, 1, 95-131.

Merrill, M. D., & Bunderson, C. V. (1981). Preliminary guidelines for employing graphics in instruction. Journal of Instructional Development, 4(4), 2-9.

Merrill, M. D., Reigeluth, C., & Faust, G. W. (1979). The instructional quality profile: A curriculum evaluation and design tool. In H. F. O'Neil, Jr. (Ed.), Procedures for instructional systems development (pp. 165-204). New York: Academic.

Merrill, M. D., Richards, R. E., Schmidt, R. V., & Wood, N. D. (1977). The instructional strategy diagnostic profile training manual. San Diego, CA: Courseware, Inc.

Murray, D. M. (1978). Internal revision: A process of discovery. In C. R. Cooper & L. Odell (Eds.), Research on composing: Points of departure (pp. 85-103). Urbana, IL: National Council of Teachers of English.

Nathenson, M. B., & Henderson, E. S. (1977). Problems and issues in developmental testing. National Society for Performance and Instruction Journal, XVI(1), 9-10.

Nathenson, M. B., & Henderson, E. S. (1980). Using student feedback to improve learning materials. London: Croom Helm.

National Assessment of Educational Progress (1977). Write / Rewrite: An assessment of revision skills; selected results from the second national assessment of writing. (Tech. Rep. No. 05-W-04). Washington, DC: US Government Printing Office. (ERIC Document Reproduction Service No. ED 141-826).

Nevo, D. (1977). A model for the utilization of formative evaluation in the process of developing instructional materials. Programmed Learning and Educational Technology, 14, 2, 127-133.

Newell, A., & Simon, H. A. (1972). Human problem solving. Englewood Cliffs, NJ: Prentice Hall.

Nichols, F. W. (1981). Generalist to specialist. Whom do I consult? National Society for Performance and Instruction Journal, XX(8), 23-24.

Nold, E. W. (1980). Revising. In C. Frederiksen & J. Dominic (Eds.), Writing: The nature, development and teaching of written communication: Vol 1 (pp 67-80). Hillsdale, NJ: Erlbaum.

O'Donnell, H. (1985). Improving textbooks - Who is responsible? Journal of Reading, 29, 3, 268-270.

Ohlsson, S., & Langley, P. (1985). Psychological evaluation of path hypotheses in cognitive diagnosis. In H. Mandl & A. Lesgold (Eds.), Learning issues for intelligent tutoring systems. New York: Springer.

Perl, S. (1979). The composing process of unskilled college writers. Research in the Teaching of English, 13, 317-336.

Pflieger, J., Chomienne, M., Bordeleau, P., & Stolovitch, H. (1979). Enquete sur l'education des documents audiovisuels dans la Province du Quebec. GERDAVE (Rapport No. 8). Montreal, Canada: Universite de Montreal, Section Technologie Educationelle.

Piaget, J. (1926). The language and thought of the child (Marjorie Gabin, Trans.). New York: Harcourt Brace.

Polanyi, M. (1962). Personal Knowledge. Chicago, IL: University of Chicago.

Rahmlow, H. (1971). Using student performance data for improving individualized instructional units. Audio Visual Communications Review, 19, 2, 169-183.

Redish, J. C. (1980). Readability. In D. B. Felker (Ed.), Document design: A review of the relevant search. Washington, DC: Document Design Project, American Institutes for Research.

Reigeluth, C. M. (1982). The elaboration theory's procedure for designing instruction. Journal of Instructional Development, 5(3), 22-32.

Reigeluth, C. M. (1983). Instructional design: What is it and why is it? In C. M. Reigeluth (Ed.), Instructional design theories and models: An overview of their current status (pp. 3-36). Hillsdale, NJ: Erlbaum.

Reigeluth, C. M., & Merrill, M. D. (1979). Classes of instructional variables. Educational Technology, 19(3), 5-24.

Reitman, W. R. (1965). Cognition and thought. New York: Wiley.

Richard, J. F. (1986). The semantics of action: Its processing as a function of the task (Rapport de Recherche No. 542). France: Institut National de Recherche en Informatique et en Automatique.

Richey, R. (1986). The theoretical and conceptual basis of instructional design. London: Kogan Page.

Robeck, M. D. (1965). A study of the revision process in programmed instruction. Unpublished master's thesis, University of California, Los Angeles.

Roberts, W. K. (1979). Some considerations for evaluating instruction using the analytical approach. Educational Technology, 19(2), 39-41.

Rohman, D. G. (1965). Pre-writing: The stage of discovery in the writing process. College Composition and Communication, 16, 106-112.

Rohman, D. G., & Wlecke, A. O. (1964). Pre-writing: The construction and application of models for concept formation in writing. U.S. Office of Education Cooperative Research Project, No. 2174. East Lansing, MI: Michigan State University.

Rosen, M. J. (1968). An experimental design for comparing the effects of instructional media programming procedures: Subjective versus objective procedures (Final Report). Palo Alto, American Institute for Research.

Saettler, R. (1968). A history of instructional technology. New York: McGraw-Hill.

Saroyan, A., & Geis, G. L. (1988). An analysis of guidelines for expert reviewers. Instructional Science, 17, 101-128.

Scriven, M. (1967). The methodology of evaluation. AERA Monograph Series on Curriculum Evaluation, 1, (39-83).

Simon, H. A. (1973). The structure of ill-structured problems. Artificial Intelligence, 4, 181-201.

Simon, H. A. (1978). Information processing theory and human problem solving. In W. K. Estes (Ed.), Handbook of learning and cognitive processes: Vol. 5 (pp. 271-295). Hillsdale, NJ: Erlbaum.

Simon, D. P., & Simon, H. A. (1978). Individual differences in solving physics problems. In R. Siegler (Ed.), Children's thinking: What develops (pp. 325-348). Hillsdale, NJ: Erlbaum.

Smillie, R. J. (1985). Design strategies for job performance aids. In T. Duffy & R. Waller (Eds.), Designing usable texts (pp. 213-243). Orlando, FL: Academic.

Smith, G. L. (1982). Revision and improvement: Making the connection. In R. A. Sudol (Ed.), Revising (pp. 132-139). Urbana, IL: ERIC.

Sommers, N. (1980). Revision strategies of student writers and experienced adult writers. College Composition and Communication, 31, 4, 378-388.

Stakenas, R. G., & Mayer, H, (1983). Formative evaluation in vocational education. Performance and Instruction Journal, 22(5), 23-36.

Stallard, C. K. (1974). An analysis of the writing behavior of good student writers. Research in the Teaching of English, 8, 206-218.

Stolovitch, H. D. (1982). Application of the intermediate technology of learner verification and revision (LVR) for adapting international instructional resources to meet local needs. National Society for Performance and Instruction Journal, XXI(7), 16-22.

Stolovitch, H. D. (1983, April). Formative evaluation research overview and update. Paper presented at the 22-nd Annual Conference of the National Society for Performance and Instruction, Detroit, MI.

Swarts, H., Flower, L. S. , & Hayes, J. R. (1984). Designing protocol studies of the writing process. In R. Beach & L. Bridwell (Eds.), New directions in composition research (pp. 53-71). New York: Guilford.

Thorndyke, P. W., & Stasz, C. (1980). Individual differences in procedures for knowledge acquisition from maps. Cognitive Psychology, 12, 137-175.

Tulley, M. A. (1985). A descriptive study of the intents of state level textbook adoption procedures. Educational Evaluation and Policy Analysis, 7(3), 289-308.

Tyler, R. W. (1941-42). General statement on evaluation. Journal of Educational Research, 35, 492-501.

Weber, R. P. (1985). Basic content analysis series: Quantitative applications in the social sciences. Bevereley Hills, CA: Sage. Winograd, T. (1972). Understanding natural language. Cognitive Psychology,3, 1-191.

Winograd, T. (1983). Language as a cognitive process: Vol. 1. New York: Addison-Wesley.

Wright, P. (1977). Presenting technical information: A survey of research findings. Instructional Science, 6, 93-134.

Wright, P. (1978). Feeding the information eaters: Suggestions for integrating pure and applied research on language comprehension. Instructional Science, 7, 249-312.

Wright, P. (1985). Editing: Policies and processes. In T. M. Duffy & R. Waller (Eds.), Designing usable text (pp. 63-96). Oralando, FL: Academic.

APPENDICES

APPENDIX I-A

DESCRIPTION OF YOUR TASK

During the course of instructional materials development, it is suggested that materials be reviewed by various experts and potential learners to enhance the effectiveness of the final product.

As a Content Expert who teaches microbiology and is knowledgeable about dental hygiene students, you are given the task of revising a section of the accompanying module which has been drafted for first year dental hygiene students in the United States. This section is the second unit and is called The Relationship of Microorganisms to Environmental Conditions, Disease and Immunity.

Revising is broadly defined as any of the changes you make excluding the total replacement of the text. For example you may want to alter the vocabulary, add or take out examples, change some of the content or presentation. Anything is acceptable.

While you are revising, you are also asked to speak your thoughts out loud so that your thought processes could be recorded. This technique is called think aloud and as you now know, it means verbalizing out loud all the thoughts, questions, comments, strategies that go through your mind while you are performing a task. This does not mean analyzing what you are doing; just amplifying your thoughts. Make sure you write in the changes you think are necessary and verbalize what you write.

It is very important that all the revisions be clearly legible. If you happen to change your mind once you have already revised a portion, simply cross it out with a single line and write in your new comments. Every comment you make is valuable to this research. **Your revised text will represent a product that is considered a final draft by you.**

Thank you again for participating in this study and

REMEMBER TO:

WRITE IN AND VERBALIZE ALL YOUR REVISIONS
THINK ALOUD AS YOU REVISE

APPENDIX I-B

DESCRIPTION OF YOUR TASK

During the course of instructional materials development, it is suggested that materials be reviewed by various experts and potential learners to enhance the effectiveness of the final product.

As an Instructional Designer, you are given the task of revising a section of the accompanying module which has been drafted for first year dental hygiene students in the United States. This section is the second unit and is called The Relationship of Microorganisms to Environmental Conditions, Disease and Immunity.

Revising is broadly defined as any of the changes you make excluding the total replacement of the text. For example you may want to alter the vocabulary, add or take out examples, change some of the content or presentation. Anything is acceptable.

While you are revising, you are also asked to speak your thoughts out loud so that your thought processes could be recorded. This technique is called think aloud and as you now know, it means verbalizing out loud all the thoughts, questions, comments, strategies that go through your mind while you are performing a task. This does not mean analyzing what you are doing; just amplify your thoughts. Make sure you write in the changes you think are necessary and verbalize what you write.

It is very important that all the revisions be clearly legible. If you happen to change your mind once you have already revised a portion, simply cross it out with a single line and write in your new comments. Every comment you make is valuable to this research. **Your revised text will represent a product that is considered a final draft by you.**

Thank you again for participating in this study and

REMEMBER TO:

WRITE IN AND VERBALIZE ALL YOUR REVISIONS
THINK ALOUD AS YOU REVISE

APPENDIX II

STIMULUS TEXT

ENVIRONMENTAL CONDITIONS, DISEASE AND IMMUNITY (1)

INTRODUCTION

Certain conditions in the environment promote growth of microorganisms while other conditions actually slow growth or kill microorganisms (2) Depending upon these conditions, the microorganisms can produce disease in humans or they can be rejected by barriers produced by the body (3)

OBJECTIVES

1 Define and spell terminology (4)

a aerobic (5)
b anaerobic (6)
c resident (7) microorganism (8)
d infection (9)
e autogenic (10) infection (11)
f environment (12)
g cross infection (13)
h virulent (14)
i immunity--natural, (15) artificial, active-(16) passive
j phagocytosis (17)
k portal of entry (18)
l immunity--natural (19)

2 List at least four types of environmental factors which affect microorganisms (20)

3 Describe the role microorganisms play in the process of infection (21)

4 List and explain the barrier the body has to infection (22)

5 List the different types of immunity and give and example of each (23)

6 List the points (portals) of entry of microorganisms (24)

7 Describe and give an example of each method of disease transmission (25)

ENVIRONMENTAL CONDITIONS

The environmental conditions which affect the growth and multiplication of microorganisms include: (26)

1 type of food available, (27)
2 moisture (28)
3 temperature (29)
4 oxygen requirements (30)
5 amount of light present (31)

It is important to realize that different microorganisms require varying environmental conditions for optimal growth. (32) These

conditions become expecially important in relation to sterilization and disinfection. (33)

RELATIONSHIP TO STERILIZATION AND DISINFECTION

It is essential to know which of these environmental conditions can be controlled to destroy harmful microorganisms and attain sterilization (34) For example, if we know a certain microorganism requires a temperature of 98 to 100 degrees Fahrenheit to live, then raising the temperature several degrees above that level may kill it (35) This is essentially what is done during the sterilization process (36)

A progrom to control harmful microorganisms in the dental office environment consists of those sterilization methods which kill the most resistant and harmful microorganisms (37)

FOOD, MOISTURE, AND TEMPERATURE

Microorganisms each have their own specific requirements for these three environmental conditions (38) Much research has been conducted to determine the combinations of these conditions necessary to maintain the life of harmful microorganisms (39) Sterilization often changes one or all of these conditions to kill the microorganism and thus make it harmless (40)

OXYGEN AND LIGHT

Oxygen (40a)

Microorganisms also differ in their need for oxygen (41) Aerobic microorganisms require oxygen in their environment in order to grow and multiply (42) Anaerobic microorganisms require an absence of oxygen in their environment in order to grow. (43) In other words, if any oxygen is present in the environment, the anaerobic microorganisms will die (44)

Light (45a)

Microorganisms also differ in the amount of light they require to grow and multiply (45) Some require a great deal of light while others actually require darkness to promote growth (46)

VIRUSES	Viruses have varying requirements for environmental conditions in which to grow and multiply. (47) Specific viruses require specific conditions (48) One thing that is essential to the growth of any virus is the presence of some living tissue. (49)

ROLE OF MICROORGANISMS IN PRODUCING DISEASE (50)

MICROORGANISMS IN THE BODY

Some microorganisms are helpful to man and some are harmful.(51) It is possible that in some quantities some microorganisms have no effect on the body, but in larger quantities may produce disease (52) The body varies in its response to the presence of microorganisms (53)

RESIDENT MICROORGANISMS

Some microorganisms are normally present within the body (54) These are known as resident microorganisms (55) The oral cavity is a good example of a part of the body where resident microorganisms may be found (56) It is important to understand that the resident microorganisms normally found in one person's mouth may be different from those microorganisms found in someone else's mouth (57) The resident microorganisms found in the mouth may not be normally present in other parts of the body (58)

TRANSIENT MICROORGANISMS

Transient microorganisms are not normally present in the body (59) They are transient, or present temporarily, in the body (60) These microorganisms may be present in the environment or be the resident microorganisms from someone else's body (61) In addition, microorganisms which are resident to one area of the body may be transient to another area of the body (62)

For example, certain microorganisms are normally found on the finger or are resident there (63) If you rub your eye with your finger, some of the resident microorganisms may be transferred to the eye, where they are not normally found (64) They would then be called transient microorganisms and in some instances could cause disease (45)

TYPES OF INFECTION

Infection When a microorganism enters the body, multiplies, and causes a reaction, this process is known as an infection. (66)

Autogenous infection An infection is known as an autogenous infection when it is caused by a resident microorganism (67) It is usually the result of resident microorganisms penetrating into an area where they do not normally reside. (68) For example, if resident

microorganisms in the mouth penetrate the oral tissues during an injection of local anesthesia and cause an infection, this would be known as an autogenous infection. (69)

Cross infection An infection is known as a cross infection when it is caused by microorganisms transferred from one person to another (70) For example, if you did not wash your hands properly (or wear gloves) and a microorganism from your hand entered a patient's mouth and caused an infection, it would be known as a cross infection (71)

Carrier of infection: A "carrier" is a person who carries or has present in the body a disease-producing microorganism, but does not exhibit observable symptoms of the disease (72) However, the carrier can transmit the disease-producing microorganism and the disease to any other individual who may be suseptible to the particular microorganism (73) For example, an individual may carry the microorganism that causes hepatitis and yet never exhibit the symptoms of the disease or every have hepatitis. (74) When the carrier comes into contact with others (s)he can transmit the actual disease to the other individuals (75)

FACTORS NECESSARY FOR THE DEVELOPMENT OF INFECTION

Several factors are necessary for the development of an infection (76)

1 The microorganism must enter the body in sufficient numbers (77)

2. The microorganism must be virulent (78) This means that it must be capable of overcoming the body's defenses and capable of destroying healthy tissue. (79)

3 The body must be suseptible to the disease. (80) In other words, if the body has developed immunity to the disease, it will not be affected by the microorganism (81)

4 The microorganism must be transmitted through the proper route, or portal of entry (82)

PORTALS OF ENTRY

For infection to occur, it is necessary for the microorganism to enter the body via an appropriate route. These points, or portals of entry, include

1. the skin, usually at the site of a cut or scrape,
2. the respiratory tract;
3. the oral cavity;
4. the nasal passages,
5. the digestive tract,
6. the genitourinary system,
7. the placenta,
8. any site where trauma occurs

A microorganism at one site may not cause a problem, but may cause an infection when it enters at the appropriate portal of entry. For example, a microorganism which causes a respiratory infection must enter through the mouth or nose to produce an infection in the lungs.

FUNCTION OF BARRIERS TO INFECTION

In order to produce a disease or infection, the microorganism must overcome various barriers on the body's surface or at the portals of entry. The function of the barriers, then, is to prevent the microorganisms from entering the body and therefore prevent disease.

CATEGORIES OF BARRIERS

The body has a number of different barriers which prevent microorganisms from producing disease. These barriers may be categorized as 1) mechanical, 2) chemical, and 3) physiologic.

MECHANICAL BARRIERS

The human body is designed in such a way that it actively and mechanically prevents a great number of microorganisms from entering it. The skin is a good example of such a mechanical barrier. The cells in the top layer of skin are "tough" and prevent microorganisms from entering the body, much like the skin of an orange protects the fruit inside. The body also has various reflexes which provide mechanical barriers to microorganisms. Examples of protective

reflexes are the blink of an eye or a sneeze (103) Both of these reflexes help to prevent microorganisms from entering the body (104)

CHEMICAL BARRIERS

The second category of barriers are chemical in nature (105) Various body fluids contain ingredients which destroy microorganisms. (106) An example of this type of barrier is the fluid known as gastric juice, found in the stomach (107) The gastric juices contain certain chemicals which help to destroy microorganisms and thus resist infection. (108)

PHYSIOLOGIC BARRIERS

Physiology is the science that deals with the function of various parts and organs of living organisms (109)

Physiologic barriers, the largest category of barriers to infection, are the natural barriers produced by normal body function (110) These body functions resist the invasion of microorganisms (111) Three examples of physiologic barriers are phagocytosis, temperature increase, and immunity (112)

PHYSIOLOGIC BARRIERS PHAGOCYTOSIS

Phagocytosis is a natural defense mechanism In this process, certain white blood cells seek out disease-producing microorganisms, surround them, and actually destroy them (113)

PHYSIOLOGIC BARRIERS TEMPERATURE INCREASE

The body also reacts to microorganisms by an increase in temperature (114) The proper temperature is needed by microorganisms in order to grow and multiply (115) When the body's temperature rises, microorganisms are no longer able to grow and are usually destroyed (116)

PHYSIOLOGIC BARRIERS IMMUNITY

The body's most complex barrier to microorganisms is known as <u>immunity</u> (117) Immunity is a complex reaction that resists invasion and disease production ability of microorganisms (118) Immunity may be broken down into the following categories. (119)

I. Natural or inherent immunity Natural immunity occurs in certain races and individuals (120)

A. Race Some races of people are resistant to certain diseases and the microorganisms that produce those diseases (121)

B. Individual Some individuals have something in their genetic structure that makes the individual naturally resistant to certain diseases (122)

II. Acquired immunity Immunity acquired after conception can be natural or artificial in nature (123)

A Natural
1. Active This immunity is acquired by the attack of a disease (124)
2. Passive This immunity is transferred from mother to child through the placenta during pregnancy (125)

B Artificial
1 Active Vaccination produces this immunity (126) By introducing specific microorganisms into the body (in small amounts) the body's immune system is activated (127)
2 Passive By injecting a specific substance into the body, immunity is produced for a short period of time (128)

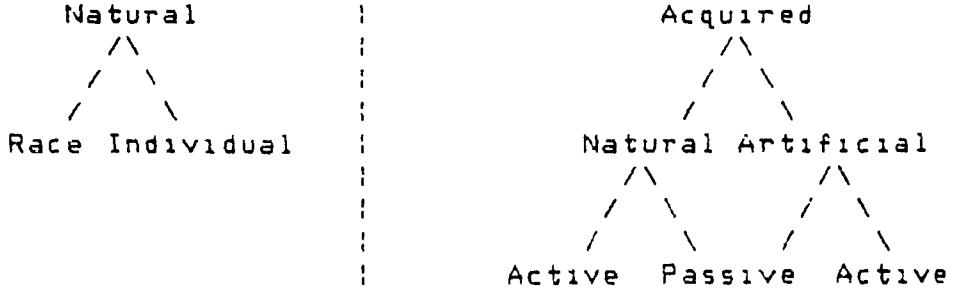

Fig. 1 Diagram of Immunity Systems (129)

INTRODUCTION

Microorganisms can be transmitted in three different ways: (131)

*direct contact (132)
*indirect contact (133)
*insect carrier (134)

OBJECTIVES

After reading this section and completing the related self-test, you will be able to describe the methods of disease transmission. (135)

DIRECT CONTACT

Direct contact is an immediate transfer of microorganisms to an appropriate portal of entry into the body. (136) An example of direct method of transmission would be touching a contaminated person. (137) "Contaminated", in this context, means that the person has the harmful microorganism. (138) Other examples of direct transmission are kissing, sneezing, coughing, talking, and sexual intercourse. (139)

INDIRECT CONTACT

Indirect methods of transmission include vehicle-borne, vector-borne, and airborne. (140)

*Vehicle-borne. The microorganism is transmitted via contaminated materials or objects (vehicles) such as toys, handkerchiefs, soiled clothing, bedding, surgical instruments, water, food, etc. (141)

*Vector-borne. This refers to simple mechanical carriers of disease such as crawling or flying insects. (142)

*Airborne. This involves aerosols (minute particles) that are suspended in the air. (143) These aerosols are transferred to a suitable portal of entry through the air. (144)

INSECT CARRIERS

Insect carriers provide a third method of transmission of microorganisms. (145) A good example of an insect carrier is a certain species of mosquito that is responsible for carrying the microorganism that causes malaria. (146)

DIFFERENCE BETWEEN INSECT VECTORS AND INSECT CARRIERS (147a)

Insects may be classified as either vectors or carriers. (147) Insect carriers are those insects that carry specific microorganisms responsible for specific diseases, while insect vectors may carry a number of different kinds of organisms due to the unclean environment in which they live. (148)

SELF-TEST 2

1. Name at least three of the five environmental conditions which affect the growth of microorganisms. (149)

A. _____

B. _____

C. _____

D. _____

E. _____

(150)

2. Microorganisms which require oxygen in their environment are called _____ (151)

3. What special environmental conditions do viruses require that other microorganisms do not require? (152)

_____

4 Microorganisms normally present in the body are known as _____ microorganisms, while those which are temporary are known as _____ microorganisms. (153)

5. When a microorganism enters the body, multiplies and causes a reaction, the process is known as a(n) _____ (154)

6. An infection caused by resident microorganisms is known as a(n) (155)

a cross infection (156)
b autogenous infection (157)
c transient infection (158)
d carrier infection (159)

7 What four factors are necessary for the development of infection? (160)

A _____
B. _____
C _____
D _____ (161)

8 Name at least five of the eight portals of entry discussed in the module. (162)

A _____ E. _____
B _____ F _____
C _____ G _____
D. _____ H. _____ (163)

9 The skin is an example of a(n) _____ barrier to infection. (164)

SELF-TEST 2

10. Which of the following is <u>not</u> considered to be a physiologic barrier? (165)

a gastric juice (166)
b pagocytosis (167)
c immunity (168)
d body temperature (169)

11. Individuals who have a resistance to disease because of their genetic structure have what type of immunity? (170)

_____

12 Vaccinations are considered to be what type of immunity? (171)

_____

13 Three metnods of disease transmission include insect carrier, _____, and _____ (172)

14 Airborne transmission of microorganisms would be an example of the _____ method of disease transmission (173)

ANSWERS TO SELF-TEST 2

---

1. A food
   B. moisture
   C temperature
   D oxygen
   E. light

   See page 12-13

2 aerobic, see page 13

3 presence of living tissue, see page 14

4 resident, transient, see page 15

5 infection, see page 15

6 b, see page 15

7 A sufficient numbers of microorganisms
   B virulent microganisms
   C susceptible person
   D proper portal of entry of microorganisms

   See page 16

8 A. skin
   B respiratory tract
   C. oral cavity
   D. nasal passages
   E. digestive tract
   F genitourinary system
   G placenta
   H. site of trauma

   See page 17

9 mechanical, see page 17

ANSWERS TO SELF-TEST 2

10. a; see page 18

11. natural or inherent; see page 19

12. acquired, artificial, active; see page 19

13. direct contact, indirect contact; see page 20

14. indirect contact, see page 20

APPENDIX III

QUESTIONS SUBJECTS FREQUENTLY ASK ABOUT TALK-ALOUD PROTOCOLS OF THE WRITING AND REVISION PROCESS*

WHAT IS A TALK-ALOUD PROTOCOL?

A protocol is a sequential recording of a person's attempt to perform a task. Developed early this century, protocol analysis is a powerful tool in educational research. In particular, the information captured in 'talk-aloud' protocols enables the educational researcher to construct detailed models of human thinking processes, and in some cases to stimulate these processes in a computer program. In short, protocols give the researcher a 'window' through which to look at otherwise invisible mental processes that occur from moment to moment. We are concerned in this short explanation with talking-aloud protocols of writers revising a text. Listed below are some more questions that subjects of protocol experiments ask.

HOW IS A TALKING-ALOUD PROTOCOL MADE?

The procedure is really very simple. The researcher will ask you to talk-aloud while you are revising or rewriting a particular document, You are to say out loud what you are thinking, You are not to worry about speaking correctly, stopping in the middle of thoughts or sentences, etc., but you should try to verbalize as *continually* as you can during the entire time you are at work. Pauses in your talk will naturally occur, but try to avoid them. If the researcher feels you are not talking often enough, he/she may prompt you.

SHOULD I TRY TO EXPLAIN HOW MY WRITING PROCESSES WORK, OR HOW I WOULD USUALLY DO THIS TASK?

Subjects who ask this question are usually trying to do the researcher's work themselves, at the same time they are revising or rewriting. You are not to describe what you "would' do, *but only what you are actually thinking about at the time you are working.* In fact, you are *not* expected to 'analyze' your writing habits or creative processes at all. You are not being asked to 'introspect', or to give an explanation or interpretation of your writing. You need only say what is on your mind at the moment. Concentrate on the task you have been given, and simply say aloud whatever occurs to you..

HOW CAN A PROTOCOL CAPTURE MY THINKING PROCESSES IF I CAN'T SAY ALOUD EVERYTHING I AM THINKING?

Of course, you will not be able to say *everything* you are thinking when you are completing even a simple writing or revision task. A portion of your thinking is lost and falls between the cracks. But almost everything you do say is valuable to the researcher. The amount of information retrieved from the talk-aloud method probably exceeds the amount to be gained by any other research method currently employed for the study of how people write and revise. Moreover, your transcribed protocol is also studied by the researcher; your finished or revised text is also studied, and compared with the talk-aloud transcript. By itself, your finished text tells the researcher very little about the *processes* you used to create the text. However, when your finished text is 'matched' with the protocol transcript, the researcher has a much more detailed picture of how your writing and revising unfolded. Again: concentrate on the task and on whatever you are conscious of as you work. Say aloud everything that comes to mind.

DOESN'T TALKING-ALOUD INTERFERE WITH MY THINKING, SO THAT I AM NOT THINKING AND WORKING AS I NORMALLY WOULD?

This question is often asked, and rightfully so. It's a very important question for researchers to deal with. At the present time, no one knows for certain if talking-aloud does interfere with your thinking during problem-solving. A lot of research is presently being conducted to find out. So far, researchers have been unable to find any strong evidence that talking-aloud interferes with thinking. Some research has even shown that, with very little practice, you can solve the same problem in the same amount of time whether you are talking aloud or not. Talking loud can also improve decision-making, and many people talk to themselves when they write anyway. The first few minutes of a protocol may feel awkward, but with a little practice this feeling will disappear and you will feel more comfortable.

SHOULD I WRITE AND TALK AT THE SAME TIME, OR ONLY BEFORE OR AFTER I WRITE SOMETHING DOWN?

You should *talk as continuously as possible,* whether you are writing or not. Sometimes you will find yourself only able to say exactly what you are writing on paper. This is perfectly fine, so long as you don't pause too long between words. If you do, the experimenter will prompt you.

SHOULD I TALK-ALOUD EVEN IF I AM JUST RE-READING WHAT I'VE WRITTEN?

Yes, you should. Avoid the temptation to mumble if, and when, you re-read your text. Even if you are skimming rapidly, and not re-reading sentences in their entirety, talk-aloud and make sure your voice is audible and clear.

---

* The above text is duplicated from:
Hayes, J. R., Flower, L., Schriver, K., Stratman, J., and Carey, L. (1985). Cognitive processes in revision. (Tech. Rep. No. 12, Appendix B, pp. 1-3). Pittsburgh: Carnegie-Mellon University, Communications Design Center.

** Ericsson and Simon (1984) have made a distinction between talk-alouds and think-alouds. They define the former as verbalization of thoughts which are encoded in verbal form in memory. If, however, the encoding is in a different form, such as visual, they use the latter term to describe the procedure. In situations when it is not clear which term describes the verbalizations better, they have used the term think-aloud. To circumvent ambiguities, in the present study, all verbalizations are referred to with the more general term, think-aloud.

APPENDIX IV

PARTICIPANT CONSENT FORM

I AGREE TO PARTICIPATE IN THE STUDY RELATED TO THE REVISION OF INSTRUCTIONAL MATERIALS. I UNDERSTAND THAT THE DATA WILL BE KEPT ANONYMOUS, EXCEPT WHEN I REQUEST INFORMATION FOR PERSONAL FEEDBACK. I UNDERSTAND THAT THE DATA WILL BE ONLY USED FOR THE PURPOSE OF RESEARCH ON IMPROVING AND REVISING INSTRUCTIONAL MATERIALS.

NAME: _____

PARTICIPANT

DATE: _____

APPENDIX V
CODING SCHEME

Evaluation Statements:

The outcome of a comparison between the currently observed state and the goal state, that is what the reviser thinks the text ought to be, is coded as an Evaluation Statement (ES). By virtue of this definition, this category represents positive and negative comments (which do not explicitly state the source of the problem), expressions of preference, judgement, internal feelings and observations, all expressed in the context of the subject's task representation.

- *I don't particularly like it/*
- *you've got a lot of white space here which is fine/*

Problem Identification:

As a specific case of Evaluation Statements, Problem Identifications (PI) contains explicit reference to an observed problem.

- *again, I find that the heading's really not well done/*
- *and right justified margin makes the question difficult to read/*

Revision Statements:

Explicit text related changes intended to transform the current state to the goal state are coded as Revision Statements (RS).

- *there should be a comma after 'disinfection'/*
- *'cross infection' should be hyphenated/*

Knowledge Statements:

Expressions of personal knowledge, including both declarative and procedural knowledge, are coded as Knowledge Statements (KS). These statements are often provided in addition to, or in lieu of, a Problem Identification, or as a reason for suggesting a particular change.

- *I don't particularly like it/ (ES)*
- *most microorganisms especially those that cause disease don't require light at all/*
- *and the orthotropes which do require light are -not important in medical microbiology/*

- *that's true only to a certain extent/ (ES)*
- *but more importantly would be the virulence of the micro-organism, how severe, how effective is it in causing disease/*
- *so it is more important to consider the property rather than the number of micro-organism/*

Text Knowledge:

Comments representing Knowledge directly acquired from the stimulus text are coded as Text Knowledge (TK).

- *um, so I can assume that cross-infection is not from your body/*
- *that's one type of chemical, but that's already in the stomach/*

Verbatim Statements:

Sections which are read aloud from the text are coded as Verbatim Statements (VS).

Text Talk:

Segments which refer to various parts of the text, but not in verbatim format, are coded as Text Talk (TX).

- *so there's a self test to test these objectives/*
- *so its gonna go into talking about those three/*

Task Talk:

Reference to an activity which is currently being undertaken or suggests the setting up of a short term goal are coded as Task Talk (TT).

- *I'm reading the introduction/*
- *let's see if they discuss this in the next section/*

Strategy Talk:

Reference to a course of action which is representative of actions normally undertaken by the subject in similar situations, but is not tied to the current task are coded as Strategy Talk (ST).

- *I usually check to see if the objectives match the text/*
- *I like to get a sense of the material first, so I'll just skim through/*

Dialogue:

Questions addressed to the experimenter, and the experimenter's subsequent responses are coded as Dialogue (D).

Boundary Markers:

An interjection which marks either a pause or a break in a thought sequence (e.g., ok, um, eh, etc.) is coded as Boundary Marker (BM).

False Starts:

Clauses which do not represent a completed thought are coded as False Starts (FS).

Unrelated Talk:

Comments which are neither related to the text nor the task are coded as Unrelated Talk (UT).

APPENDIX VI

EXCERPTS FROM PROTOCOLS

EXCERPTS FROM CONENT EXPERT (I)*

A: -ok /BM
-I'm reading the definition of my task /TT
-I'm ready to read the instructions /TT
and verbalize
-I don't like the word verbalize /ES
-ok /BM
-(introduction .....body) /VS
-uh /BM
-I would not put /FS
-unless this is a very broad introduction to the entire eh-eh-this entire thing, I would not put immunity and the environment in the same spot /RS
-ok /BM
-now /TX
-(objectives ... autogenous) /VS
-yes, /ES
-that's fine/ ES
-(....) /VS
-that seems quite an appropriate terminology although I don't use autogenic /ES
-oh yes, I could use it /ES
-ok /BM
-(list at least four types.....) /VS
-I am not quite sure what is meant / ES
-but I suppose /FS
-I don't see what they mean, /ES
-oh I see / ES
-ok /BM
-(moisture, temperature) /VS
-fine /ES
-ok / BM
-(describe the role ... infection) (second reading: role in the -process of infection)
-(role....) /VS
-that /FS
-uh - / BM
-fine, /BM
-I don't sort of like the word 'role' /ES
-but I I would accept it /ES
-its ok /ES
-I think I don't know what is meant by role /ES
-(list and explain....) /VS
-fine / ES
-(List ....) / VS
-fine /ES
-(.....) / VS
-fine ES
-I would list first that /RS
-and then speak about immunity /RS

* Dots indicate sections read from the text

B: -(skin....., the placenta) /VS
-why is placenta included? / PI
-I suppose the - after the,- its very specific
-its one of the very specific things included in this list of very general things /ES
-ok / BM
-alright, that's eh /FS
-one might leave it /ES
-one could remove it /RS
-(any site where..) /VS
-because there could also be other specific things that could be included if one were to include placenta
-ok /BM
-(a M .... cause an injection .. portal of entry) /VS
-fine /ES
-it was already stated before /ES
-(for example ... lungs) /VS
-that's fine as an example /ES
-(function of barriers ... infection) /VS
-I wonder why barriers are always,are barrier /FS
-ok /BM
-(in order to produce a disease .. surface) /VS
-where am I? /TT
-ok /BM
-(the function of barriers)/VS
-uhum / ES
-(categories mechanical, .. physiological) /VS
-I I would skip be.. being sort of impatient /TT
-and just continue down /TT
-(mechanical barriers) (skin) /VS
-mucus surface should be in there somewhere /RS
-(.. tough and prevent M from .. orange ... to M) /VS
-and I don't quite understand this sentence /ES
-(examples of ... both of these reflexes help ... body) /VS
-well alright /ES
-fine /ES
-I would put FS
-however I would include mucus surface with mechanical barrier /RS
-yeah I would

C: -and I have a test
-Ok/(number 1.....) /VS
-fine /ES
-They just have to read what they were provided with and they should have no difficulty. /ES
-(Number 2....) /VS
-that's fine /ES
-(number three...) /VS
-a good question /ES
-(number four ..) /VS
-ok they call them residents /VS
-(...) /VS
-ok / ES
-(....) /VS

-ok /ES
-cross infection. /VS
-they presented it this way so I suppose they could ask the question this way /ES
-but cross infection and carrier infection is very much the same thing /PI
-so I wouldn't use that /RS
-(...) /VS
-so they have listed four /ES
-ok /BM
-there are all sorts of portals of entry /KS
-there's the skin etc. /KS
-(...) /VS
-ok /BM
-I suppose they mean the temperature /ES
-no I don't like that question /ES
-(physiologic barrier) /VS
-its temperature /KS
-its, they listed "gastric juices, phagosytosis and immunity" /TK
-I guess body temperature probably interferes /KS
-ok /BM
-although that is also physiological /KS
-THE BODY TEMPERATURE
- (...) /VS
-they are speaking here about racial immunity and such /TX
-(....) /VS
-ok /BM
-they want the artificial and active /KS
-(three methods of disease..) /VS
-oh /BM
-they want exactly the three which they listed /KS
-alright /ES
-(reading twice) /VS
-that is the droplet transmission /KS
-I suppose the aerosols /KS
-no its / FS
-I don't like that /ES

EXCERPTS FROM CONTENT EXPERT (II)

A:
-We'll begin with this module here /TT
-the spelling of sterilization is wrong /PI
-there is and 'i' missing after the 'r' /PI
-and there should be a comma after 'disinfection' /RS
-so the title should read (MICROORGANISMS RELATED TO STERILIZATION AND DISINFECTION) with an 'i' /RS
-comma /RS
-(environmental conditions, disease and immunity) /VS
-the word rejected is not a good choice /PI
-it could probably be modified by another word like

inhibited /RS
-the section on objectives/TX (define and spell terminology) /VS
-I don't understand why they have used verbs instead of nouns /ES
-I think that it's a bad kind of heading for this group of words /PI
-they should probably be or it should probably be definitions and spelling terminology, nouns instead of verbs /RS
-and those words there /TX
-'cross infection' may be hyphenated/RS
-and I don't know why there is a redundancy between 'i' and '1' /PI
-(immunity.....natural) /VS
-there should not a repetition because '1' is already included in the item 'i' /RS
-(reading # 2, 3, and 4) /VS
-it is probably to PREVENT infection /RS
-(reading # 5) /VS
-the (and) should read AN/RS -(reading # 6, 7) /VS
-now is this for the microorganism or is it for the host? /ES
-I suppose this is the type of food available to the microorganism rather than to the host /KS
-(reading # 2,3,4 and 5) /VS
-why is light important? /PI
-I don't know of any microorganism that is affected by light /KS
-(it is ...growth) /VS
-now amount of light is probably not important /KS
-(expecially) should be ESPECIALLY /RS

B:
-(food, moisture.....) /VS
-and that is a lot of nonsense because sterilization doesn't need to take into account any of these things /PI
-sterilization should probably consider the type of microorganism present /KS
-the volume of material to sterilize /KS
-but not really the conditions of growth of these organisms/KS -sterilization is not based on moisture and temperature for the growth of the microorganisms /KS
-it is a drastic and absolute process that kills them /KS
-(oxygen and light.... multiply) /VS
-that's true /ES
-(aerobic.....to grow) /VS
-that's true and yet /ES
-in other words if any oxygen spreads in the environment of anaerobic microorganisms, some of them might survive without growing and stay put /KS
-they are tolerant of anaerobic condition, of aerobic conditions rather /KS
-and light, we said, that is not an important factor for pathogenic microorganisms /KS
-there are no microorganisms that are infective that require light to grow /KS

-(viruses.....multiply) /VS
-and that is like describing viruses as if they were extra cellular/KS-that is not true /PI
-all viruses require would be the presence of a living cell /KS
-and this living cell has to be receptive to invasion by the viruses /KS
-that is the only requirement /KS

C: -let's go through it and see /TT
-(microorganisms.....) /VS
-its a different subject /ES
-this has nothing to do with disease transmission /PI
-(microorganism..... carrier) /VS
-ok,/these are different methods of transmission /TX
-(objective.....such as toys) /VS
-and so forth /TX
-one word that people would prefer to use would be formite-borne/-BY MEANS OF F-O-R-M-I-T-E /ES
-(vector.....)VS-now, this is not /FS
-well, I'll come back to that in a minute /TT
-(airborne...air) /VS
-ok, that's quite correct /ES
-I'll go to vector-borne there because the next section we talk about insect carriers /TT
-(insect carriers.....microorganisms.....malaria)/VS
-now this kind of insect transmission is different from that other fly because here the micro. goes through a sexual cycle or develops within the insect while in the vector-borne it is strictly a mechanical carrier/KS-that is it is not a biological carrier /KS
-(difference between) /VS
-oh, here we are/TT(insect vectors .....live) /VS
-that's one way of putting it /ES
-yeah, we prefer to, at least I prefer to see the insect carriers defined as those in which a specific micro. grows within the insect tissue rather than carrying a specific kind of micro./KS-so that's about it. /TT

EXCERPTS FROM CONTENT EXPERT (III)

A:
-(reading title)/VS
-I'm going to read the introduction out loud /TT
-(reading introduction) /VS
-ok /BM
-I've read the introduction/TT-and it seems to be very self-explanatory and very simply stated /ES
-I think first year students will be able to understand that /ES
-I don't think I'd change anything for now/ ES
-objectives of this /FS

-I'm going to read them out loud now /TT
-(reading first objective) /VS
-so those, that's the first set of objectives /TX
-that's quite a lot for this section /PI
-I hope that they're going to remember all those things /PI
-that's ten, twelve different very new eh terms in this sect, section here/eh uh, not having heard any,, a lot of these words before, I'm not sure that these students will be able to understand it all /PI
-now, second objective is: /TX
-(reading # 2) /VS
-ok /ES
-four types of environmental factors is not that many /ES
-that sounds ok /ES
-(reading # 3 /VS
-that's pretty straightforward /ES
-eh there are many different roles that microorganisms can play in the process of infection /KS
-and so I'm assuming that they are going to talk about one general type of role that could sort of overlap into all the different microorganisms /ES

B:
-(portals of entry.....) /VS
-there's eight different sites there /TX
-ok /BM
-(micro.....lungs) /VS
-ok /BM
-eh, it says here about (may cause and infection when it enters at the appropriate portal of entry) /TX
-and I guess I was thinking it'd be appropriate for the micro., but inappropriate for the human body /PI
-so, maybe changing that word to something else would be less confusing /KS
-like appropriate usually means beneficial /KS
-and its not beneficial to the human body for these micro. to be entering into them /PI
-and I'm taking the point of view from the human side not the micro. side /KS
-ok /BM
-so I might change that word, /RS
-or put in commas after the word /RS
-(function of barriers to infection.....disease) /VS
-ok /BM
-that's quite simply stated /ES
-I think I understand all that /ES
-and even just the title makes sense /ES
-I like that idea /ES
-(categories of barrier...) /VS
-right /BM
-I'm just thinking for myself/ TT mechanical, the eyelashes, the skin, nails and so on and so forth /KS
-chemical against the different, when we have tears, the chemical reaction of destroying the bacteria that might enter in our eyes /KS

-and the physiological that'd be blood cells and antibodies that are being produced and chemical too, that's be the acid in our stomach that would destroy a lot of bacteria that we eat with the food

C:
-(reading question # 8)
-ok, that's a good thinking question too and strictly memory recall
-lots of good questions
-number nine you give away one of them by giving this question
-(the skin... barrier to infection)
-ok, so that's a good question too
-most of these questions are recall memory type of questions
-(reading # 10)
-good, that's a good question cause one needs to know what all these, how all these different things react to and know whether or not it is a physiological barrier or is it a chemical barrier or what
-ok
-I like that question
-that's a good question
-you have to think about that one -(reading # 11)
-ok
-good question
-(reading # 12)
-right, what it makes me think of is what I was having a hard time with before
-(reading # 13)
-ok
-that's a good question
-(reading 14)
-that's a good question too
-maybe there's a way of not making so long sp, such a long space between "an example of the"
-so I'm putting a line between those words
-so that it would work out a little bit easier for -and here's all the correct answers

EXCERPTS FROM INSTRUCTIONAL DESIGNER (I)

A:
-ok/
-(M related...disinfection...immunity..introduction)/
-I see its organized by headings, objectives and then content headings/
-ok/
-and this is a module/
-(introduction ...body), (objectives..define and spell)/
-vocabulary words that are supposed to be used in the module

I assume to be able to define and spell/ ES
-(four types of environmental factors which affect M- the role of M -the barrier the body has to infection- natural barrier-list the different types of immunity and give examples -list the points of entry -describe and give an example of each method of disease transmission)/ VS
-so there's a self test to test these objectives/ TX
-self test two (checking)/ ES
-combination of short answer, multiple choice and their answers /ES
-ok/ BM
-(environmental condition which affect the growth of.. multiplication of M includes)/ VS
-which is objective two/ ES
-(type of food available, moisture temperature-oxygen requirements-amount of light)/ VS
-ok/ BM
-as far as reading, I find the right column justification difficult because of the spaces it creates between the words/ PI
-ok/ BM
-the headings don't seem to correlate specifically with the objectives/ PI
-so I assume that first section must have to do with the second objective/ ES
-let's see/ TT
-(relationship to sterilization ...disinfection .... Microorganism/ VS
-ok/ BM
-so these are ranged in order of the five environmental conditions/ ES
-I would find it more helpful if the (oxygen) headings of the , of the content sections were the same as the headings in the list of five under environmental conditions /RS
-seems like they're not even the same. /PI
-page fifteen starts the third objective /ES
-(M..... disease..The Body varies in its response to the presence of M, Resident M, transient M, types of infection, factors necessary for the development of infection, portals of entry, functions of barriers)/VS
-I think some kind of structural overview would be helpful for this section /RS
-It's hard to know how these headings relate to each other/PI
-and how the information is to be given /PI

B:
-and let's see - if the objectives match the requirements/TT
-(list, describe, list and explain) /VS
-list and explain goes beyond rote recall, I would say/KS
-(list the different types of immunity, list the points, describe and give an example) /VS
-so the questions are I guess within keeping of the kinds of things that the objectives call for except the possible exception of objective four /ES
-I don't see any questions that ask for the explanation of

anything / PI
-ok /BM
-I find this organization much too open to / PI
-and given the nature of the questions that they're all rote would be very easy to simply match questions as you go along /ES
-and if I'm correct, seems that the questions are also in fairly close, similar order to the objectives which are the same order as the text /ES
-yeah /ES
-I think so /ES
-so it really lends itself to that approach for the person going through the module which again is not an aid to any sort of long term retention sense / PI
-people are not being asked to manipulate any ideas but simply to recall rotely / PI
-that I'm not sure as a result of this module that the retention would be good / PI
-Alright /BM
-(Name at least three of the ..M) /VS
-ok /BM
-(type of food available, moisture, temperature, oxygen requirements) / VS
-ok / BM
-they want three of the five but five spaces are given / ES
-(of light present) / VS
-so that's a problem / PI
-if three of the five are asked for, there should only be three spaces / RS
-(M.. called) /VS
-big space between are / PI
-and again right justified margin makes this question difficult to read / PI

C:
-so I find that I was able to complete this exam with only one out of fourteen wrong (which is a good mark) without reading the text /ES
-and I dare say if somebody asked me this information tomorrow, I would have, I think I would remember that there are different kinds of infection / TK
-there are different kinds of immunity / TK
-that there are different kinds of bodies that carry infection / TK
-eh /BM
-I think that's about all I can remember /ES
-so if it were up to me to revise this unit, I would first of all look at what kind of information I wanted to be rotely remembered / RS
-and why and how it linked to information that needed to be retained at a higher level
-and I would create some kind of meaningful structure for the person going through this module / RS
-so they could see how the informations presented relates to itself
eh

-I would ask people to do some sort of prediction structure at the beginning to bring out the information they already know and therefore have some link /RS
-I find that the thing that's difficult about this information especially (may not be for somebody who's more familiar with content but for myself) is the fact that I was /PI
-nothing was established so that I could link the newer information.
-I'm learning to do it /TT
-I find the quiz wholly inadequate because its its too easy to take /PI
-and as I said before by just going through the material without reading it
-I think that I would /FS
-everything is of equal importance supposedly, visually by inference to the overall content /PI
-and in fact some of these categories break down /RS
-so that some of these categories can be under each other /RS
-I find that difficult. /PI

EXCERPTS FROM INSTRUCTIONAL DESIGNER (II)

A:
-Okay, well we might as well just take it from the top /TT
-In normal situation would be to familiarize myself with the material, so I would skim the material briefly first /ST
-(reading) Okay, then /SM
-luckily this has objectives listed, already so we have some, /ES
-uh first one define and spell terminology /TX
-uh /BM
-(.....), etcetera, etcetera, (..... transmission) /VS
-and then it goes on to talk about environmental conditions /TX
-I'm going to look back, basically to the title /TT
-okay, /BM
-(Microbiology related... immunity) /VS
-I will reread the introduction /TT
-then um /FS
-I will then again take a quick look at the objev, /TT
objectives to make sure that they all refer to the title and the introduction
-Which I assume they would be /TT
-um immunity /TX
-(.....) /VS
-so the other thing that I would tend to look for in a upper level class like this would be is, a graduate level class, /FS
-Um there are some, certainly some low level cognitive objectives as well as some upper level cognitive /ES
-and I basically look only on those levels --Low being rote, and upper level being whatever, going up the taxonomy /TT
-Okay then I will continue to uh /TT
-well actually what I would do next, would ah, be to go to

the back and take a look at the self-test, and just to make sure that um, the self-test refers to the objectives.
-On page twenty-one /ST
-Okay, so the self-test, okay, um, /TX
-question one (.....) /VS
-okay /BM
-so I've got to define and probably spell this terminology as well, /TX
-so these are the definitions on that /TX
-six (.....) /VS
-five would be (reading)/ VS
-question six (...) /VS
-number thirteen (airborne transmissions) /VS
-so that's related to those /ES
-and then the answers /TX
-well it seems to, to relate fairly straightforward to it/ES
-um, then what I would do, would /ES
-I'm continuing on my brief uh, runthrough of it, just to get, eh, if I can an idea of the way that the learning is structured /ST
-so, following through the objectives, /TT
-after that, um, there would be, um, they're giving a definition here /TX
-the first thing environmental conditions, /TX
-what they mean by environmental conditions /TX
-and list different aspects of, of environmental conditions /TX
-The relationship to sterilization and disinfection
-so it's more information /ES
-you know these are basically giving me the uh, some definitions and terminology to support the, uh, what I have to know, for the most part / ES
-And viruses /TX
-So those are all basically definitions /ES
-Then it goes on to talk about microorganisms and the body/ TX
-Again these are statements and, of definitions /ES
-Uh, then listing factors necessary for development of uh, development of infection / TX
-and several factors involved there which relate distinctly to the objectives /ES
-And then portals of entry /TX

B:
-One of the problems here that I'm, that I'm seeing so far is that, no doubt this was written by a subject matter expert. / PI
-It doesn't, it doesn't tell a story, /PI
-okay, I mean, I, I, I like, you gotta have a beginning a middle, and an end /RS
-and it's gotto follow logically / RS
-and the material is presented in the kind of format where somebody sits down, okay, this is first, this is second, that's third, / PI
-and they're listed as isolated segments without the transitionary pieces between them, /PI
-as well as visuals and everything else, /PI

-that would make it a bit more understandable. /RS
-I mean I understand what they're talking about, /ES
*but just to provide ease of access to the learner.
-Okay so we got the function of the barriers, /TX
-no problem there, /ES
-categories we learned about, /TX
-mechanical barriers one of the three barriers a mechanical barrier, /TX
-I mean that, that's you know, that kind, /TX
-(human body is... entering the body) /TX
-Um, it's funny, /ES
-I, I don't know if I would have called a blink or a sneeze a mechanical barrier, / PI
-but I guess, I guess it is. /ES
-I'd like to see what they say about physiologic. /TT
-Okay, (Chemical... stomach) /RS
-I guess it uh, makes it unpleasant for the microorganisms to survive /ES
-(Certain chemicals... infection) /RS
-Okay, so, um, that's one type of chemical, but that's already in the stomach /TK
-I'm just wondering if there are other chemical barriers elsewhere in the body /ES

C:
-Um, the content, let's, let's take in order here, /TT
-so I think that's one thing, to, to find out what's really important. /RS
-The objectives are listed here, are relative okay, /ES
-the course does not necessarily relate to the objectives /PI
-and the um, final test does not relate totally to the objectives. /PI
-I mean, yes, I mean, if you look at em, the words are the same, but it's not getting at the behaviors that the objectives want to get at. /PI
-Um, there's also no practice within the um, uh, course material itself on the, on the objectives. /PI
-Now you could say that self-test is a practice, but in the subject matter, which may or may not be totally new stuff to dental hygienists, /PI
-so I mean, they probably are not familiar with this, with this, uh, with the information, /PI
-and therefore it should be presented in a, a more bite size, /RS
more appropriate for the learner approach, RJ/ with presentation of information, with examples, RJ/ uh, all these quote unquote good things that one does in designing an instructional program. /RS
-Which this one is lacking. /PI
-Um, and if you did that, if you spent uh, you know, a week with it, I think you can come up with something that's a little bit better. /RS
-Again. I think you do need visuals with it, /RS
-I think that would help quite a bit. RS
-I mean, we're dealing with things that um, uh, they're

abstract to a certain degree, yet it's concrete because we're talking about your body, /KS
-but, you ask fifteen people to draw a microorganisms and they'll probably draw something that look like an amoeba or something like that. KS
-And that may not be totally accurate, KS
-and ah, I think they need to know, I mean, one thing, how many microorganisms does it take, uh, to cause a disease, you know, /KS
-they say you gotta have sufficient number, and yet before they say it's, you know, they multiply, so it could be one ever that starts it. /PI
-So, um, there are a lot of questions that are unanswered by the material, /PI
-and I think even if you gave to a subject matter expert they would find some uh, factual errors. /ES
-Anyway, I mean, so subject matter expert can pull it apart. /ES
-I think from instructional point of view, uh, we could go through it piece by piece, and look at it, but I, I, /T
-I think from an overall point of view, this was not developed in a instructional design format. /PI
-Somebody has a concept of objectives, and a concept of a test situation relating to the objectives, but everything in between is basically content. /PI
-Okay /BM
-And content doesn't work as an instructional, uh, as a way to carry instruction, you can't carry instruction by content alone, /KS
-you know what you have to do with it. /T
-So, um, this is not done appropriately. /PI

EXCERPT FROM INSTRUCTIONAL DESIGNER (III)

A: BM
-Okay/ I'm doing what I normally do, which is to just get a feel for things, /ST
-I don't like to start reading something till I sort of know what I have ahead of me. /ST
-And I see that it's broken down into Introduction objectives, Environmental conditions, and so forth. /TX
-It looks like the type is pretty dense, on this page and it's making me uncomfortable, /PI
-I'll see what we can do about that later, /T
-I like it a little bit airier. /ES
-Okay, looks like there are very few diagrams, it's mostly words, /ES
-and there it looks like there's a subtest at the back of the module. /TX
-Does not appear, ES
-and an answer guide for that. /TX
-Does not appear that there's much opportunity for practice throughout, /PI
-looks like it's mostly content /ES
-and not any breaks for application. /PI

-but we'll see when I get into it,
-Okay, so I've now got a general feel for what's going on.
-Okay, Introduction (certain conditions....by the body).
-Okay that's kind of like background information.
-Generally when I look at something like this, I see a model in my head that's says that part of the introduction should include a rationale for why, why this would be important to the learner.
-So I'm just going to make a note here that, that says rationale for learner.
-The background statements are fine but, what's in it for me,
-in other words, how does it apply to my job is what I'm asking myself now.
-Okay, now I'm going to scan the objectives,
-and (Define and Spell Terminology...)
-Oh God, I don't even know what these things mean -There's a typo there,
-I'M just going to remove the
-D (List the ....)
-Okay, I don't know whether these objectives are stated in terms of the learner or in terms of the instruction.
-I might suggest adding here something like AT THE END OF THIS MODULE YOU WILL BE ABLE TO,
-and then each of the points following, hold me responsible, or accountable,
-I like to see that in the front,
-although, maybe it's done in the first module,
-and if so, maybe it's not necessary to continue it through out,
-maybe it's done for the overall course,
-I don't know,
-so I'm going to make an assumption here that it's not
-and I want to be consistent if it is.
-Okay, I like the way the objectives have action verbs, there's no know, or appreciate or whatever,
-and that's one of things that I look for right away,
-so I'm seeing hard action verbs that tell me precisely what, I as a learner would have to be able to do,
-and so I like the way that is, listing, describing, defining, spelling, so forth.
-Okay, the objectives seem to move from sort of simple to complex,
-and their sequence and they seem to go from more cognitive things to more application orientations,
-so I have no problem with the sequence, at least at first pass.
-Okay, I'm going to move on to I think what is content at this point,
-uncomfortable with the format here in the sense that I would like to know that now I'm jumping into the actual instructional material

B:
-Okay, now this format again,

-why am I getting stars, /PI
-I don't like stars, /ES
-I mean come on let's be consistent here, /RS
-I'm gonna do the 1, 2, 3. /RS
-If we use 1, 2, 3, before, why not use 1, 2, 3.
-I think what happened is somebody took a bunch of different modules that other people had developed and put them together and said isn't this wonderful. /KS
-I'm surprised they're not even on, they are different typewriters. /PI
-Okay, well that explains it, /ES
-somebody just sort of patched it together, which typically happens, /KS
-we grab some content wherever we can find it, /KS
-and we don't worry about whether or not it's parallel construction, or whatever. /KS
-This person was real happy about to use lots of dividing lines, too, /PI
-so it's at least consistent in that respect. /ES
-(Microorganisms... carrier) /VS
-Okay (objectives, after reading... self-test) /VS
-Oh, there's a self-test for this section? /CS
-For this section or this module?
-I think this is misleading. /PI
-I don't know what they're referring to, /ES
-so I'm going to put a question mark by section and completing (YOU WILL BE ABLE TO DESCRIBE THE METHODS OF DISEASE TRANSMISSION.) /RS
-Is it to name or list and describe? /ES
-I don't know, because if I go back to my objective it says (Describe and give an example of each method of disease transmission). /PI
-So I think back there I'm going to say maybe they should (LIST, COMMA, DESCRIBE AND GIVE AN EXAMPLE OF EACH METHOD DISEASE TRANSMISSION. /RS
-Unless, in the test I'm going to be required to, they're gonna give it to me
-and I'm going to be asked to describe it.
-Well, which reminds me, I haven't been a very good girl about this, /TT
-I should've spent more time looking at the test to see what in fact is going to be used for evaluation. /TT
-Even though I'm getting close to the end, it's not too soon to catch myself and say what are they, what were they trying to evaluate. /TT
-Let's see if there's anything parallel between these objectives and the evaluation. /TT
-So I'm skipping over to page twenty-one now, because now I'm concerned what's going here. /TT
-I want to see the nature of evaluation. /TT
-That doesn't mean that we can't change the test, /TT
-but let's see what they had intended here, for purposes of comparison. /TT
-(Name at least three of the five... isms.) /KS
-well that's real interesting, since this one says, (List)

the objective says (list at least four) and the question says name at least three. /ES
-So it's definitely not parallel. /PI

C:
-Okay now I'm done with the content /TT
-and my reaction again is, I need to have a self-test here, for this third section. /RS
-And the self-test would be again, it would include terms from /FS
-And the self-test would be again, it would include terms from objective one, that were new, /RS
-oh, there were none, /ES
-but I'm going to encourage that there be some included if they're important enough to be taught /RS
-then they ought to be tested. /RS
-And also objective number seven, which is describing the examples. /RS
-So I'm gonna need a self-test that's gonna cover the third section objectives which are some terms, which I would include, and also the seven. /RS
-Now, I'm still going to turn over the page, /TT
-well no I'm not, /TT
-I'm gonna use some blank paper here, /TT
-cause I sort of scrunched that self-test in, /TT
-and I'm also going to suggest that before the self-test the big self-test for the whole module, that we have some sort of conclusion, or summary, /RS
-by way of the objectives we told them what they were going to learn. /ES
-The objectives served as our advanced organizer, /ES
-and at the end we should tell what they learned, /RS
-so we've told them what they're going to learn, /ES
-we tell them, we've told them what they've learned /ES
-and then I think we should prepare them for the uh, /ES
-so it should be in two parts, One a conclusion or summary, -and two a preparation for the self-test. /RS
-Ah, it doesn't have to be a long preparation,
-maybe we could say to them in a just a moment you're going to be taking a self-test, which will help you determine whether or not you've acquired all the skills and knowledges that this module was intended to do.
-You've already had an opportunity to check your progress along the way. /RS
-Maybe what we could do is call this a self-test /RS
-and call the other one a progress check. /RS
-Or call this a module review, /RS
-and call the other one a self-test or progress test, or whatever. /RS
-But in some way, we have to differentiate the self-test from the interim checks. /RS
-DIFFERENTIATE FROM INTERIM TESTS, or interim progress checks. /RS
-Okay, I would like to see that because this has been a big thing for somebody to bite off, /RS

APPENDIX VII-A

Table 1

Frequency of Categories: Raw Scores

	Content Expert			Instructional Designer		
	I	II	III	I	II	III
Evaluation	92(7)*	25(2)	131(9)	41(2)	78	240(5)
Problem Id.	20	22(4)	41(6)	25	80(2)	107(9)
Revision	42	22	62(14)	14	48(6)	139(43)
Knowledge	41	60	57	5	50	32
Text Know.	7	0	8	17	28	17
Verbatim	97	53	97	84	82	74
Text Talk	20	5	35	18	41	44
Task Talk	15	5	27	8	22	127(8)
Strategy T.	0	1	0	4	10	14
Dialogue	2	0	0	9	67	15
Boundary M.	50	0	119	32	22	41
False Start	21	1	15	4	21	7
Unrelated T.	1	0	4	2	19	25
TOTAL	408	194	596	263	568	882

* Numbers in parenthesis indicate repetitions. These amounts have been included in the main column.

APPENDIX VII-B

Table 2

Frequency of Categories: Percentage Scores

	Content Expert			Instructional Designer		
	I	II	III	I	II	III
Evaluation	22.5	12.8	21.9	15.5	13.7	27.2
Problem Ident.	4.9	11.3	6.8	9.5	14.0	12.1
Revision	10.2	11.3	10.4	5.3	8.4	15.7
Knowledge	10.0	30.9	9.5	1.9	8.8	3.6
Text Knowledge	1.7	0.0	1.3	6.4	4.9	1.9
Verbatim	23.7	27.3	16.2	31.9	14.4	8.3
Text Talk	4.9	2.5	5.8	6.8	7.2	4.9
Task Talk	3.6	2.5	4.5	3.0	3.8	14.3
Strategy Talk	0.0	0.5	0.0	1.5	1.7	1.5
Dialogue	0.4	0.0	0.0	3.4	11.7	1.7
Boundary Mr.	12.2	0.0	19.9	12.1	3.8	4.6
False Start	5.1	0.5	2.5	1.5	3.6	0.7
Unrelated Talk	0.2	0.0	0.6	0.7	3.3	2.8
TOTAL	408	194	596	263	568	882