THE PRIORITIES ESTABLISHED AMONG DATA SOURCES WHEN INSTRUCTIONAL DESIGNERS REVISE WRITTEN MATERIALS

by

Cathrine Anthony Le Maistre

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Educational Psychology

Department of Educational and Counselling Psychology

McGill University

Montreal

March 1994

© Cathrine Anthony Le Maistre

ACKNOWLEDGMENTS

The term "acknowledgment" is not strong enough to represent the debt of gratitude I owe to the people who have helped shape this dissertation.

I thank the members of my committee for the help they have given and the variety of input that they have provided. Dr Robert Bracewell's expertise is evident in the design and analysis of the study, less evident, although appreciated, in the support he gave at appropriate times. Dr George Geis introduced me to the systematic design of instruction, and has been available whenever I needed his counsel and his view of the big picture. Dr Lynn McAlpine could always be relied on to ask the right questions and to press for as much clarity and precision as I could provide. My supervisor, Dr Cynthia Weston, has given me time that she often could not afford to give, and I shall always be grateful for her advice, friendship and collegiality.

I thank the members of the Formative Evaluation Research Team, for teaching me the value of working as a member of a group, and for demonstrating team spirit in its true sense. Thank you Tino Bordonaro, Nicole Duy, Alanna Israeloff, Tim Rahilly, and Diana Tremblay, all of whose contributions are noted in this work.

This research has been supported in part by a grant from the Social Studies and Humanities Research Council, whose help is acknowledged, with thanks.

Finally, this would never have been finished without the support and encouragement of my family. I thank them for their forebearance and affection. i

ABSTRACT

The objective of the study was to investigate the priorities among data sources established by instructional designers when they were given formative evaluation data from a variety of sources and asked to make revisions to instructional materials. Two modules of printed instructional material were revised by each of eight instructional designers under three treatment conditions using a counterbalanced design. The treatments were the feedback condition (feedback data from learners, subject-matter experts and audience experts, or no external feedback data); the module used; the order of administration (whether with feedback first or without feedback first). Verbalizations of all eight subjects were collected by a think-aloud procedure, then coded to reflect the problem-solving nature of the revision process. Segments relating to the revisions suggested by the subjects formed the dependent measures for analysis. Results showed that subjects were significantly more likely to incorporate their own knowledge into revisions than to use the feedback data they were given. Further analysis showed that if feedback data were used, subjects attended to learner comments in preference to any other data, and that when they made reference to feedback data from external sources, they were more likely to accept these data than to reject them.

ü

RESUME

Cette étude visait à examiner les priorités que les concepteurs de matériel pédagogique établissent entre des sources de données lorsqu'on leur fournit des données d'évaluation formative provenant de diverses sources et qu'on leur demande de réviser de matériels didactiques. Deux modules de matériels didactiques imprimés ont été révisés par chacun des huit concepteurs de matériel pédagogique répartis entre trois protocoles, le tout selon un plan d'étude équilibré. Les protocoles se distinguaient par le type de rétroinformation (rétro-information provenant des apprenants, des experts en la matière enseignée et des experts en auditoire, ou aucune rétro-information extérieure), par le module utilisé, et par l'ordre d'administration (avec ou sans rétro-information préalable). Les propos des huit sujets ont été recueillis selon une procédure de reflexion à voix haute, puis codés afin de refléter la nature "résolutive" du processus de révision. Les segments se rapportant aux révisions proposées par les sujets constituent les mesures dépendantes aux fins de l'analyse. Les résultats révèlent que les sujets étaient beaucoup plus portés à intégrer leurs propres connaissances aux révisions qu'à utiliser les données de rétroinformation qui leur étaient fournies. L'analyse plus poussée a révélé que lorsque les données de rétro-information étaient utilisées, les sujets portaient de préférence attention aux remarques des apprenants qu'a toute autre donnée et que, lorsqu'ils faisaient référence aux données de rétro-information provenant de sources externes, ils étaient davantage portés à accepter ces données qu'à les rejeter.

iii

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	i
ABSTRACT	ü
RESUME	iii
TABLE OF CONTENTS	iv
LIST OF FIGURES	vii
LIST OF TABLES	viii
CHAPTER 1: INTRODUCTION	1
Research questions	3
Summary	5 5
	5
CHAPTER 2: LITERATURE REVIEW	6
Revision as a problem-solving act	7
Revision in composition	13
Use of the think-aloud technique	16
The lack of research-based revision heuristics	19
Textbook prescriptions for revision	25
Summary	29
CHAPTER 3: DESIGN AND METHOD	31
Design	31
Method	33
Phase 1: Collection of Feedback Data	35
Selection of Stimulus Material	35
Obtaining feedback data on stimulus material	37
Learner feedback: test data.	38
Learner feedback: oral data	40
Expert feedback: oral data	40
Phase 2: Revision of Modules	43
Subjects Used as Revisers	43
Solicitation of subjects	44
Procedure Outling of procedure	45
Outline of procedure Materials for asthering muicien data	45 46
Materials for gathering revision data Equipment	40 54
Role of researcher: during the revision task.	54
Role of researcher: after the revision task.	54
Role of coresearcher	55
Data Analysis	55
Transcription	55
Segmenting	-55
Coding	56
Development of coding scheme	56
Coding scheme	57
Reliability and stability	71
Reduction of the number of codes obtained from the data sets	72
Other sources of data for analysis	74

iv

Written comments Retrospective debriefing	v 74 75
Summary	75
	77
CHAPTER 4: RESULTS Description of Subjects	78
Qualifications	78
Preferred Reference Material	79
Affiliations and Conferences	80
Familiarity with Materials and with Audience	81
Performance of Research Task	81
Task Definition	81
Subjects' Ability with the Think-aloud	85
Mean Time on Task for each Session	85
Overview of Subjects' Revision Procedures	86
Results of segmenting the transcribed think-aloud data	89
Grouping of Revision Codes	91
Ouestion 1: How Much Personal Knowledge Do Revisers Incorporate, if	
They Are Provided With Feedback Data From Experts and Learners? <u>Ouestion 2</u> : What Attention Will Instructional Designers Give to Each Data Source If They Are Provided With Verbal Comments From Subject- matter experts and Audience Experts, Verbal Data and Test Data From	92
Learners, And Asked to Revise Print Materials?	96
Acceptance or Rejection of Feedback Data	101
<u>Ouestion 3</u> : Do Revisers Provided With External Feedback Incorporate As Many of Their Own Revision Suggestions as Revisers Provided With no External Feedback?	106
<u>Ouestion 4</u> : How Do Revisers Establish Priorities Among Data Sources?	100
Use of Feedback Data in the Current Task	108
Preference for Subjects' Own Input	108
Opinions of the Expert Feedback Data	110
Opinions of the Learner Feedback Data	111
Approach to Contradictory Data	113
Reporting of Subjects' Usual Practice	114
<u>Ouestion 5</u> : When They Revise Materials, Do Instructional Designers Follow	114
the Prescriptions of the Standard Model of Instructional Design?	116
Summary	120
CHAPTER 5: DISCUSSION	122
Overview of Study	122
Question 1: How Much Personal Knowledge Do Revisers Incorporate, if	
They Are Provided With Feedback Data From Experts and Learners? <u>Ouestion 2</u> : What Attention Will Instructional Designers Give to Each Data Source If They Are Provided With Verbal Comments From Subject- matter experts and Audience Experts. Verbal Data and Test Data	123
matter experts and Audience Experts, Verbal Data and Test Data From Learners, And Asked to Revise Print Materials?	125
<u>Question 3</u> : Do Revisers Provided With External Feedback Incorporate As Many of Their Own Revision Suggestions as Revisers Provided With	123
no External Feedback?	128
<u>Question 4</u> : How Do Revisers Establish Priorities Among Data Sources?	129

Use of Expert Feedback	vi 130							
Use of Learner Feedback: Oral Data	131							
Use of Learner Feedback: Test Data	131							
Use of Contradictory Data	132							
<u>Question 5</u> : When They Revise Materials, Do Instructional Designers Follow								
the Prescriptions of the Standard Model of Instructional Design?	134							
Revision as Problem Solving, and the Use of the Think-aloud Technique	137							
Summary	140							
Specific contributions of this study								
Theoretical Significance	141							
Practical Significance	143							
Limitations	144							
Recommendations for further research	146							
REFERENCES	148							
APPENDICES	159							
	107							
CERTIFICATE OF ETHICAL ACCEPTABILITY	263							

vii

Table	Page
1. Summary of Research Plan	33
 Percentage of Learners Who Answered Items Correctly for the Pretest a Posttest 	and 39
3. Summary of Materials Packages for Each Condition, Labelled as in	
Appendix B	47
4. Textbooks Used by Subjects During Training	79
5. Textbooks Still Consulted by Subjects	80
6. Time On Task For Each Pair of Conditions	86
7. Summary of Coded Segments for Each Subject	90
 Number and Percentage of Revision Statements Linked to Reviser's In and to All Other Feedback Data Sources for Each Subject 	iput 94
Comparison of the Number of Revision Statements Linked to the Revis Input with Input from All Other Sources	ser's 94
10. Number and Percentage of Revisions Linked to Each Source of Feedba Data	ack 97
11. Revisions Linked to Each Source of Feedback Data as a Percentage of of Revisions Linked to External Sources	Total 98
12. Significant Differences in Revision Statements Linked to the Feedback Sources	Data 100
13. Revision Statements Initiated or Corroborated by Feedback Data	103
14. Revision Statements Despite or Rejecting Feedback Data	104
15. Subjects' Estimate of Their Attention to Each Source of Feedback Data	u 106
 Comparison of Number of Revisions Linked to Revisers' Own Input i Feedback Conditions 	n Both 107
17. Data Sources Usually Available to Subjects and Their Attention to The	m 115
18. Revisers' Actions in Terms of the Standard Model of Instructional Des	sign 117

LIST OF TABLES

LIST OF FIGURES

Figure	Page
1. Model of the problem space	9
2. Formative evaluation as review plus revision	15
3. Revision as a convergence between composition and instructional de	sign 16
4. Attention to revision in instructional design textbooks	26
5. Summary of method	34
6. Comparison of reviewers for Module 1 and Module 2	42
7. Classification of coded segments to isolate revision segments	72
8. Comparison of the number of revision statements, and theoretically r	
segments with all coded segments	91
9. Revision codes linked to reviser's input or to external feedback	93
10. Number of revision statements linked to reviser's input and to all oth	her
feedback data sources for each subject	95
11. Revision codes linked to each source	96
12. Percentage of revisions linked to each data source for each subject	99
13. Effect of module on revisions based on subject-matter expert comme	ents and on
test scores	101
14. Revision codes reflecting reviser's acceptance or rejection of feedbac	ck data 102

viii

CHAPTER 1: INTRODUCTION

Instructional design consists of the systematic process of developing instruction. One of the best established models of instructional design (Dick & Carey, 1985, 1990) identifies the following components in the system: (a) identifying an instructional goal; (b) conducting an instructional analysis; (c) identifying entry behaviors and characteristics; (d) writing performance objectives; (e) developing criterion-referenced test items; (f) developing an instructional strategy; (g) developing and selecting instruction; (h) designing and conducting the formative evaluation; (i) revising instruction; (j) conducting summative evaluation.

The components of the model investigated in this study were formative evaluation and revision. Although materials developers have always made modifications to their work before using it with learners, the term formative evaluation was coined by Scriven in 1967. In the context of instructional design, formative evaluation is taken to mean the collection of data on instructional materials and the use of these data to revise the materials to improve their effectiveness.

The revision stage of formative evaluation is one of the least studied steps in the systematic design of instruction, although research suggests that improved learning outcomes are a benefit of revising instructional materials. The literature supports the use of feedback from a variety of sources, whether from learners (e.g., Henderson and Nathenson, 1977), or from experts (e.g., Nevo, 1985) or from a combination of both (e.g., Weston, 1987). Geis (1987) identified a number of different experts who may act as reviewers and provide revision suggestions, including subject-matter experts, instructional design experts and audience experts, such as teachers. Research has shown that there is consistency among the types of revisions made by revisers from the same discipline when they are not given learner or expert review data (Saroyan, 1989), that revisers pay attention to student data (Duy, 1990), and also that experts bring tacit knowledge to the revision task

(Kandaswamy, 1980). However little work has been done to examine the decisions taken by revisers when they choose among data sources when making revisions.

Most models of instructional design follow a four-stage sequence for collecting data on draft materials (see, for example, Dick & Carey, 1985, 1990):

1. Review by experts, usually subject-matter experts who provide information regarding the accuracy and currency of the materials; teachers or trainers who assess the materials for instructional appropriateness and suitability for the intended audience; and instructional designers who assess aspects of design and presentation;

2. Feedback from individual learners;

3. Feedback from small groups of learners;

4. Field testing with a large group of learners.

Learners may provide oral or written information on the weaknesses of the materials; they invariably provide test data. Revision is usually recommended after each step in the evaluation process. The degree of revision falls on a spectrum ranging from doing nothing, because the materials are already in the best possible form, to rejecting the whole package and starting the instructional design process again.

This four-stage process of collecting data and using them in revision is the standard model of formative evaluation included in the training of instructional designers. Textbooks instruct practitioners to use data collected during the review of materials but contain limited advice as to how to use such data. Although instructional designers are told to incorporate feedback data into revisions, the research literature does not indicate what process they actually follow in practice. If instructional designers do use feedback, what components of these data do they use? For what purpose do they use them? How do revisers deal with data which conflict with each other, or with their own intuition? If there is evidence that they do not use data from a particular source, this would put into question the injunction to collect such data. The need to validate the requirement for the expensive and time-consuming process of data collection is of pragmatic importance to practitioners.

The practice of instructional design is still in the process of developing a theoretical base (Richey, 1986), although it has long been viewed as a problem-solving process (Wager, 1993; Waldron, 1973) and this may prove to be a useful paradigm. Problem solving has formed the basis for the study of revision in composition (Hayes, 1989), with revision being viewed as an "ill-structured problem" (Simon, 1973). A comparison of revision as defined in the literature on composition with revision as a component in the systematic design of instruction shows that they are both examples of problem solving. Acceptance of a problem-solving model permits the development of a coding scheme based in problem-solving. It also allows the application of procedures such as the think-aloud technique (Ericsson & Simon, 1980), which have proved useful in investigating problem solving.

This study investigated the actions of professional instructional designers while they were revising. This observation of the practice of instructional designers was guided by the following questions.

Research Questions

Instructional designers are trained to incorporate feedback from experts and learners (external data) in revising instructional materials, and there is evidence to suggest that they also incorporate their own knowledge (internal data) when they revise. This study sought to investigate (a) the actions of professional instructional designers when they were asked to revise a module of self-instructional text with and without feedback data, and (b) the choices the instructional designers made among data sources. To illuminate the actions of instructional designers when they revise and to investigate the choices they make among external and internal data sources, the following general research question was proposed:

What priorities are established among data sources when experienced instructional designers revise written materials?

This question was broken down further and operationalized as follows:

1. How much personal knowledge will revisers incorporate, if they are provided with verbal data from subject-matter experts and audience experts, verbal comments from learners, and test data from learners?

This question was investigated by providing revisers with such data collected on draft materials, and instructing them to revise the materials, while thinking aloud.

2. What attention will instructional designers give to each data source, if they are provided with verbal data from subject-matter experts and audience experts, verbal comments from learners, and test data from learners, and asked to revise print materials?

This question was investigated by providing revisers with such data collected on draft materials, and instructing them to revise the materials. Actual performance on the revision task was also compared with revisers' estimates of the attention they had paid to each data source.

3. Do revisers provided with external feedback incorporate as many of their own revision suggestions as revisers provided with no external feedback?

Revisers were also asked to revise draft materials without being provided with any expert or learner feedback data, so that the only input came from their own knowledge.

4. How do experienced revisers establish priorities among data sources?

Comments made by revisers during the think-aloud process and explanations given during a follow-up retrospective interview were analyzed for the reasons proposed by revisers for their choices among all possible data sources.

5. When they revise materials, do instructional designers follow the standard model presented during their training?

Comments made by revisers during the retrospective interview were analyzed to provide a demographic profile of the subjects and to describe their assessment of their usual practice.

Summary

In summary, the present study elaborates three themes: (a) the application of a broader theory base than has traditionally been applied to instructional design; namely, problem-solving theory, the study of composition and the consequent use of the research tools of these other domains; (b) the gap in research into the actual practice of professional instructional designers when they revise; and (c) the link between the prescriptions given to instructional designers during their training and their subsequent practice. The literature related to each of these themes will be examined in Chapter 2. This will be followed by a description of the design and method of the study in Chapter 3. Chapter 4 will report the results obtained and the analyses performed, and Chapter 5 will discuss these results, draw some conclusions, and make recommendations for further research.

CHAPTER 2: LITERATURE REVIEW

Instructional design comes out of a behaviorist tradition, characterized by a focus on observable events and performances. This has been--and continues to be--useful for pragmatic application of instructional design models in the development of instructional materials. However, the movement in educational psychology from behaviorism to cognitivism is reflected in journal articles dealing with the influences of cognitivism on instructional design (e.g., Clark, 1992; Winn, 1990). Further, if internal behaviors are to be researched empirically, the behaviorist paradigm is less useful than one which allows covert processes to be examined. Cognitive science has provided both a valuable theoretical framework and a methodology. Kember & Murphy (1990) expressed the need for a "black box" in the traditional systems approach to instructional design to accommodate learner characteristics. It would appear that reviser behavior also needs a black box: Given the procedures of formative evaluation prescribed during their training, what do revisers actually do?

This review of literature has four principal components:

1. The theoretical framework of the study, which will be further broken down into revision as a problem-solving act, and the place of revision in formative evaluation and in composition;

2. The use of the think-aloud as an appropriate technique from the theory base of human problem solving for researching the actions of revisers;

3. The lack of research in the specific area of reviser performance and the consequent lack of revision heuristics;

4. The revision prescriptions provided in textbooks of instructional design.

The chapter will end with a summary of the research discussed under these headings.

Revision as a Problem-solving Act

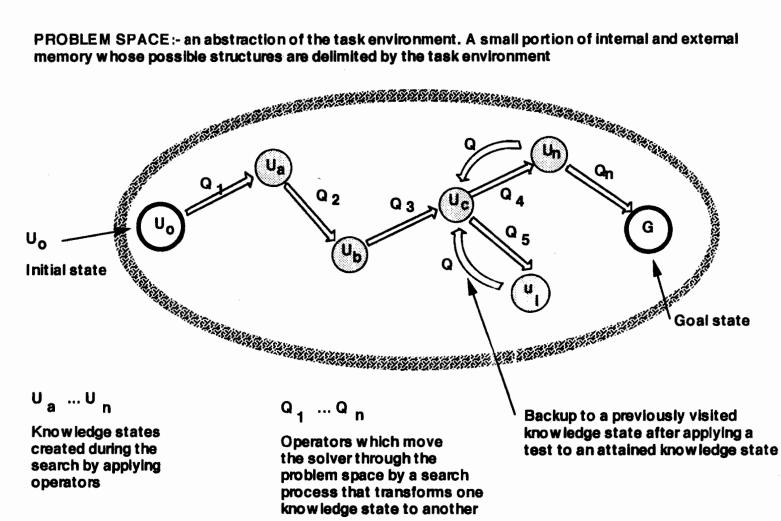
The process of designing instruction has lacked a firm foundation in theory (Foshay & Moller, 1992; Mitchell, 1989), despite attempts to establish such a foundation (Richey, 1986). One theoretical base that is beginning to seem fruitful (Saroyan, 1989) is human information processing theory, which proposes that systematic development of instruction is a problem-solving activity (Rowland, 1993; Waldron, 1973) If instructional design is considered to be "a systematic process of problem-solving" (Wager, 1993, p. 11), its sub-components, including revision, must also be regarded as problem-solving, by extension. Many definitions of problem solving exist (e.g., Hayes, 1989), some of them specific to the content domains (e.g., Polyà, 1947), but the most useful, because it applies across domains, seems to be that derived from Newell and Simon's (1972) work, "Human Problem Solving."

Newell and Simon defined problem solving as a search within a problem space "which humans have or develop when they engage in goal-oriented activity" (Newell, 1980, p. 696). They envisage the problem solver as existing in an environment which contains a task which he is motivated to accomplish (p. 55). Once within this task environment, the solver establishes the existence of a problem, by determining that there is a gap between the initial state and the goal state (p. 789, 790). The problem solver then builds a representation of the problem by setting up a problem space consisting of a set of nodes, or knowledge states. The nodes represent what the solver knows at different stages of the solution process (Simon & Lea, 1974, p. 108) and include the initial state and the goal state. A solver who is naive in the context of the problem will have few nodes in the problem space, while a sophisticated solver will have more nodes, or postulates, for the problem at hand. Whether the solver is naive or sophisticated, establishing the problem space is the most crucial step for many problems (Hayes & Simon, 1974), and the one on which most effort is directed (Simon, 1973). When the problem space has been established, the solver searches this space by applying operators, actions which transform

one knowledge state into another, and thus generates a path through the problem space (Anderson, 1993). If the operators have been correctly chosen, this process moves the solver closer to the goal state. If the problem domain is familiar to the solver, less search is involved because the solver has a repertoire of operators and knows which ones to apply, through "search control knowledge" (Newell, 1980, p. 701). It is more economical for the solver to call up a previously used problem space from long term memory; this is characteristic of "professionals" who rely more on experience and training than on explicit instructions (Simon & Hayes, 1976). The problem space model is represented graphically in Figure 1.

The design of instruction fits this model of problem solving. The first step in planning instruction is often to identify whether the situation is susceptible to an instructional solution--is there a problem? If a problem is identified in this way, probably by a needs assessment, the designer will next represent the problem, or construct a problem space, by conducting an instructional analysis to assess the learners, the context, and other constraints. Problem resolution is effected by developing instructional strategies and materials; i.e., selecting and applying operators within an instructional context. It is characteristic of the systematic nature of instructional design that solution evaluation, in the form of formative evaluation, takes place throughout the process and feeds back into the system at each stage.

The revision step of instructional design is also an example of problem solving (Yelon, 1974), and can be described by the four steps of the human problem-solving model. In this case, the problem identification would consist of recognizing that there is a flaw in the materials. This may be identified by expert review, by test data from learners, or by the reviser. When revisers set themselves the goal of improving the materials, they are constructing a problem space, or, if the problem is a familiar one, evoking a previously used space. Suggesting and incorporating alternatives would comprise the



PROBLEM SOLVING: Searching the problem space to find paths from initial state to goal state

application of operators. Solution evaluation would be done by the reviser, or by resubmitting the revised product to expert review or to learner verification.

If the task environment and the goal state are clearly identified, and if the solution is susceptible to a test, the problem is called a well structured problem (Simon, 1973). Examples might include classic chess problems, such as those investigated by de Groot (1965). If the task environment and goal state are less categorical, and if, as is often the case, the solver is required to identify the problem, and to define both the problem space and the goal state, as well as to find the path, the problem is ill-defined (Chi, Glaser, & Farr, 1988). Reitman (1965) characterized an ill-defined problem as one which does not have a generally agreed-upon answer, nor a high degree of agreement within the community of solvers. Simon's (1973) definition of an ill-defined problem was one "whose structure is lacking in some respect" (p. 181), whether in definition, solution process, constraints, or solution. In such a case, the problem is declared solved by the application of a "stop rule" (Voss & Post, 1988, p. 281). It is typical that the problem solver builds his own representation of an ill-structured problem, both at the beginning of the process, and as the solution proceeds (Simon & Hayes, 1976). Hayes (1989) described the different "gap-filling" decisions (p. 30) which are necessary to solve an illdefined problem, with the result that different solutions are to be expected.

Variation in the structure of any one of the components (definition, solution process, constraints, or solution) affects the degree of definition of a problem. Thus a problem may exist anywhere on a continuum from completely defined to completely illdefined. Revision of instructional materials may be placed at the ill-defined end of the problem continuum, because revisers' decisions whether or not to revise take place in a problem space established by the revisers, not for them. Whether or not they are provided with feedback on the materials, the revisers must form their own definition of the problem to be solved and decide what action to take. Further, although there should be general agreement that a problem is well-solved if improved learning outcomes occur, there is

usually a variety of possible solutions. The "answer" to the problem is indeterminate, and also subject to judgement by the reviser (see, for example, Duchastel, 1990; Voss & Post, 1988). The stop rule for the revision process may be when learners are able to reach a preset performance criterion (Geis, Burt ,& Weston, 1984), or often, in practice, when time or money are exhausted. It would be difficult to imagine a definitive test for a perfect revision.

If revision is considered to be an ill-defined problem, it follows that no algorithm is possible to prescribe revision procedures to practitioners, since an algorithm, if correctly applied, guarantees a single solution. At best, revisers can be expected to develop empirically a collection of heuristics to effect the process. Indeed, Duchastel (1990) maintained that instructional design is so knowledge intensive that it cannot usefully be reduced to a "procedural formalism" (p. 441).

In recent years, research in problem-solving has moved from situations which might be described as puzzle problems, or, in VanLehn's terms "knowledge-lean" problems (VanLehn, 1989, p 528) and towards "knowledge-rich" problems--more complex situations which make demands of the solver's prior knowledge and which involve a more complex problem-space. The problems confronting a reviser may be considered as ill-defined and knowledge-rich problems, since the reviser must recognize that the existing text is problematic, establish a goal state which would improve the text, and then apply knowledge to reach the goal he has defined. Nathenson and Henderson (1980) described revision as "an essentially creative process" which involves "a substantial subjective element" (p. 133), and Smith and Wedman (1988) characterized revision as "one of the most 'arty' aspects of instructional design" (p. 15). Dick and Carey (1991) used the metaphor of detective to describe the reviser.

All this suggests that the actions of the reviser are complex applications of both knowledge and experience. Such actions are demanding of both the reviser's knowledge of the process, and of decision-making skills, which may be based more on analogical reasoning than on rules (Duchastel, 1990). Shanteau (1992) reported evidence that experienced performers in a variety of domains have poor decision-making skills, and that decisions are often based on irrelevant information. He suggested that these inaccurate decisions result from the application of heuristics rather than rules. This is probably a consequence of the fact that the problems Shanteau reported are ill-defined problems. In an instructional design context, Romiszowski (1986) claimed that "One cannot escape the ultimately heuristic nature of the decision-making process involved in instructional design" (p. 433). Yet instructional design is a systematic process, and models of instructional design are rule-based. This suggests a mismatch between the process of instructional design and the type of problem to be solved by the instructional design process. If this coherence between the problem and the process is lacking, questions arise about the actual procedures followed by instructional designers.

In summary, the problem-solving steps derived from Newell and Simon's problem space model are:

1) recognition of the existence of a problem

2) evocation or construction of a problem space

3) selection and application of operators

4) evaluation of the new knowledge state attained.

These steps are taken to be homomorphic with steps in revision. They are assumed to be iterative, as is the revision phase of formative evaluation, and will be used as the basis of the coding scheme which will be described in Chapter 3. Further, revision is taken to be an ill-defined problem and therefore not susceptible to an algorithm. This raises a question regarding the behavior of revisers trained to follow an algorithmic process: if no algorithm is possible for revision, how do revisers decide what to do?

Revision in Composition

Another domain in which human problem-solving theory has been productive is that of composition. Revision of instructional materials, whether in print or in any other medium, is comparable with revision in composition. The writer has a goal in mind, identifies a feature of the writing that does not meet the specifications of his goal, and considers alternative forms. Solution evaluation would require a decision based on feedback from a reader or on the experience of the writer.

A point of contrast is that revisions of instructional materials may be subjected to more empirical tests than revisions of writing. Instructional designers may ask questions related to the teaching and learning effectiveness of the revised materials, which may be repeatedly revised until no further improvement in learning is observed. While this is theoretically possible, little research has been done to prove whether or not it is the case. In practice, decisions regarding how many cycles of revision to conduct are more commonly made on pragmatic grounds of cost and time involved.

The major contribution to the study of composition as a problem-solving activity has been made by the team at Carnegie-Mellon University (Flower, Carey, & Hayes, 1985; Flower & Hayes, 1980; Flower, Hayes, Carey, Schriver, & Stratman, 1986). They have defined revision in writing as a problem-solving activity and further, as an ill-structured task which is easier for experienced writers because they have "a well-learned, 'off the shelf' script to guide them" (Carey, Flower, Hayes, Schriver, & Haas, 1989). They do not regard this script as constituting an algorithm. In their description of the writing process, Flower and Hayes (1980) defined reviewing as being the sum of reading and editing, where editing has the purpose of correcting the text. But the same team described a group of revisers among their subjects as making changes that would preserve as much of the text and the content of the text as possible (Flower, Carey, & Hayes, 1985). Elsewhere they wrote that "*Revise* refers to the process of deciding that one will use diagnostic information to fix text problems" (Hayes, Flower, Schriver, Stratman, & Carey, 1982, p.

229). This definition of revision has proved useful for this study, although it has been modified since Hayes et al. did not specify the source of the diagnostic information, whether from reader feedback or from the reviser.

Fitzgerald (1987) summarizes various conceptions of revision as "making any changes at any point in the writing process" (p. 484), which is too general to provide guidelines for the subjects of an experiment such as the one included in this study, while Debert's (1979) advice to instructional designers to: "add something, simplify something, change something" (p. 20) or Nathenson and Henderson's (1980) "add, delete, move, modify" (p. 137) seem too simplistic. Faigley and Witte (1981) developed a taxonomy which separated revisions into "surface changes" and "meaning changes" (pp. 402, 403).

In the present study, revision is defined as any changes, short of rewriting, that the reviser deems necessary to improve the materials. Changes included in revision may be substantive or at a surface level. At a substantive level, the reviser may restructure the materials by adding, deleting or moving content so as to make the material better for instructional purposes. Thus, the reviser's knowledge must include instructional psychology and materials design. <u>Review</u> is defined as the assessment of draft instructional materials, with the intention of subsequently improving them. The review may be performed by learners, experts of different sorts, or by the revisers. Data may be in the form of test scores, responses to questionnaires or reactionnaires, or oral comments and may reflect cognitive or affective concerns. When the review is performed by the reviser, it may be explicit, and conducted as a separate stage of formative evaluation, or it may be implicit, and apparent only by its impact on the revisions produced. This study defines <u>editing</u> as a sub-set of revision, constituting the superficial level changes made during revision. Editorial changes are concerned with such issues as errors of typography, grammar, punctuation and spelling, where these errors do not impact the instructional quality of the materials. These definitions have the advantage of separating the review process (or problem identification phase) of formative evaluation from the revision (or

problem solution phase). By separating revision from review, it also admits the possibility of revision taking place without previous review by agents other than the reviser.

This is summarized in Figure 2.

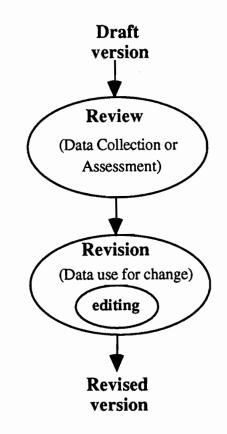


Figure 2. Formative evaluation as review plus revision.

There is enough similarity among the definitions of review, revision and editing used in this study and those used in the composition literature to recognize common ground between these two different domains. At first glance, instructional design and composition may appear to have little in common. The design of instruction is considered to be systematic, based on highly structured procedures, while writing appears to be more intuitive and unsystematic. Yet the two processes have much in common. Both fit the same model of problem-solving and both contain a clearly defined revision step which also fits the same model. Both benefit from the same research methodology and each may benefit from research carried out in the domain of the other. The process of revision may be seen as a convergence between composition and instructional design, and both composition and instructional design fall within the problem-solving paradigm, which is, itself, subsumed under human information processing theory. (see Figure 3.)

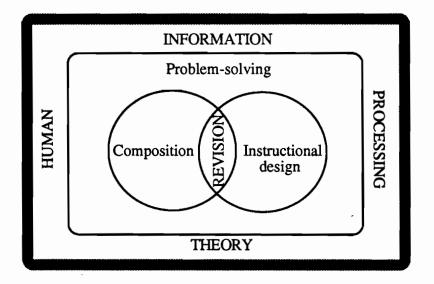


Figure 3. Revision as a convergence between composition and instructional design.

Use of the Think-aloud Technique

One of the advantages of applying the problem-solving model to both composition and formative evaluation is that it allows the methodologies of problem solving, which have been validated in composition research, to be applied in research on formative evaluation. Specifically, the use of the think-aloud technique and subsequent analysis of the protocol produced by thinking aloud has proved to be useful to track cognitive activity in both these domains (e.g., Flower, Carey, & Hayes, 1985; Markle, 1989; Rowland, 1992).

The procedures of protocol analysis are founded on a theoretical framework of human information-processing theory (Ericsson & Simon, 1980) and also on the attempts of Newell & Simon (1972) to develop a computer simulation based on how people think. The think-aloud procedure is a well-validated method of probing the cognitive processes of a subject. There is a large body of evidence that the analysis of the protocol derived when subjects think aloud allows internal processes to be inferred from overt verbal behavior (e.g., Smagorinsky, 1989; Steinberg, 1986), since subjects "verbalize the information they attend to while generating the answer" (Ericsson & Simon, 1993, p xiii). It appears that verbalizing does not interfere with cognitive function although, depending on the complexity of the task, it may slightly decrease the speed of task performance (Ericsson & Simon, 1980), and can be tiring for the subject. It is clear that the analysis of verbal protocols has allowed insights into cognitive processes which would otherwise be hidden, and that it is an effective tool for examining the processes of formative evaluation (Backler, 1990; McAlpine, 1987; Saroyan, 1989; Smith & Wedman, 1988).

Newell and Simon propose that although evocation of a previously used problem space involves long term memory activity, search through a problem space takes place in short term memory. In a think-aloud, subjects are instructed to think out loud, as they might talk to themselves if they were alone, but not to try to explain their thoughts. They are asked to read aloud any text they read or write, and are told not to worry about hesitations, ungrammatical speech, nor the triviality of what they are saying. The intention is that their verbal productions will reflect "information currently being attended to" (Ericsson & Simon, 1980, p. 225) or "such as we have of something we are in the act of doing" (Polanyi, 1959, p. 12). Thus, subjects' verbalizations are restricted as far as possible to low level processing and to reporting current thoughts and activities. These current thoughts and activities occur in short term memory, as does search through the problem space. The think-aloud protocol should follow the subject's path through the problem space and provide a trace of this path.

In addressing the reliability of such a trace, Genest and Turk (1981) reported evidence that for low level processing, such as reporting what is in "focal awareness" (Polanyi, 1958, p. 56), verbalizing has little effect on thinking. They also found that in

subsequent analysis of protocols, lower level responses are more reliable because they are less reactive; that is, as long as the subject stays in the lower levels of performance, there is no editing, nor does the subject apply a selection mechanism to what is verbalized. In fact, Ericsson and Simon (1980) based their model of protocol analysis on verbalization of "only information in focal attention" (p. 235). Higher level responses have less reliability, and demand more interpretation. If the subject is carrying out very high level thought, verbal production shuts down completely (Ericsson & Simon, 1980; Genest & Turk, 1981), and prompting is used as a strategy for maintaining verbal continuity. The experimenter should take care not to prompt too early because prompting gives rise to retrospective reporting of thoughts, which are subject to editing and to theorizing. These are part of the role of the experimenter (Newell & Simon, 1972), rather than the subject.

In the second edition of their book, Ericsson and Simon (1993) summarized the current description of think-aloud protocols as "sequences of thought" representing "at least a subset of the thoughts heeded while completing a task" (p. xxxv). Think-aloud protocols do not describe details of the heeded information, nor the reasons for its retrieval from long-term memory.

Ericsson and Simon (1993) cite evidence in support of using think-alouds to study expert performance, "to assess the cognitive processes, knowledge and acquired mechanisms that mediate the superior performance of experts" (p. xxxviii). They acknowledge that the performance of experts may have become automatized, so that the verbal protocols may be providing an incomplete trace rather than a detailed record of the cognitive processes. But expert performance is qualitatively different from novice performance in that it involves the strategic accessing and retrieval of the expert's knowledge base. This leads Ericsson and Simon to claim that verbal protocols obtained from experts are more revealing about processes than those obtained from novices, since the processes themselves are more revealing. The extensive knowledge base of experienced performers make the think-aloud an especially appropriate research tool to investigate their behavior, since the knowledge of an experienced performer can interfere with other methods of eliciting data, such as retrospective reporting or explanation. Macdonald-Ross (1989) cited tacit knowledge as the reason why performers, especially "master performers" (p. 147), may carry out a task but be unable to give reasons for their actions. Genest and Turk (1981) claimed that "conscious thought about an overlearned behavior can interfere with its execution" (p. 239). Tiemann and Markle (1984) wrote that "people who competently make many fine distinctions may not be consciously aware of the basis (the attributes) which guide their judgements" (p 27). This lack of awareness may produce gaps in the protocol of an experienced performer drawing from long term memory during a task, so that an incomplete record may result (Ericsson & Simon, 1980). Even so, it is a better solution than having the performer try to explain an action which may be so familiar that he is unable to explain it.

Even though thinking aloud is demanding, the reports may contain irrelevancies and will certainly have gaps; no collection method yields all possible data. The protocols produced by thinking aloud are rich and dense data sources, and are accepted as a valuable research tool for tracking the thoughts passing through short term memory. For these reasons, it was decided that a reviser's decision to use one data source over another would be better described if the subject performed a think-aloud than if the revised material were inspected after the fact, and that the revisers' strategies might be inferred accurately from their observed actions and verbalizations.

A more extended description of protocol analysis and the think-aloud procedure, and of the recommendations followed in the present study may be found in Ericsson and Simon (1984 and 1993).

The Lack of Research-based Revision Heuristics

In a 1983 presentation, Stolovitch reviewed much of the research associated with formative evaluation and concluded with the thought that the step between obtaining feedback data from learners and revision is difficult to control, partly because of reviser effects. This seems surprising since it might be expected that instructional designers have followed a similar training in the systematic approach, and would follow the same procedures; that is, to collect feedback from experts and learners, obtain test scores, and incorporate them into revisions. The concept of reviser effects is less surprising when examination of the literature shows the lack of explicit revision heuristics and provides instances of when revisers are assumed to be using their own intuition. Some key studies of formative evaluation conducted over the past 25 years will be examined, and their application to the present study will be summarized.

Early examples of formative evaluation were applied to the development of programmed learning (e.g., Baker, 1970). More recently, examples of formative evaluation have come from distance learning (e.g., Willis, 1993), multimedia (Reynolds & Ehrlich, 1992), or computer software (Piette & Smith, 1991). Whatever the delivery vehicle, there seems to be little change in the directive: collect data and use them to revise. There is little evidence of the "set of empirically or theoretically based ideas" that Yelon (1974) claimed formed the basis of the "intuition, insight, and a good dose of common sense" (p. 292) that revisers might call into play.

The literature is replete with examples of authors commenting on the lack of a revision heuristic (e.g., Briggs & Wager, 1981; Dick, 1980; Engler, 1976; Geis, 1987; Kandaswamy, 1980; Henderson & Nathenson, 1977b; Thiagarajan, 1978). According to Smith and Wedman (1988), "There are few guidelines, much less principles, to guide revisions" (p. 15). McCormick (1976) deplored the lack of a revision model as a defect in revision theory, and one which "renders much of the evaluation efforts useless" (p. 208). It seems that such a heuristic has not been forthcoming because of the lack of empirical

information about what revisers actually do when they revise. Although there is general agreement on the superiority of materials revised according to formative evaluation procedures, there is far less agreement on what processes cause this effect. Medley-Mark and Weston (1988) suggested that the revisers' interpretation of data and their subjective abilities will affect the revisions they make. Saroyan & Geis (1988) noted that it is unclear whether the improved text is produced by "intuition and personal skill, or a systematic use of heuristics or procedures" (p 103). Two reasons for this may be that reviser behavior has not been studied to any great degree, possibly because of the absence of effective research tools, and that existing studies of reviser behavior have shown mixed results.

In general, studies have either focused on the type of data needed to inform revisions, or on descriptions of the improvement in learning outcomes produced by various feedback sources. Nevo's (1985) study was one of the few to describe revisions based on expert feedback but does not describe how the revisions were conducted. More commonly, studies describe the influence of learner data on revision.

Rosen (1968) showed that results based on student data were better than those based on intuition which, in turn, were better than unrevised materials. But while Rosen's study attempted to control for the effects introduced by using different revisers, his definition of revision as a process of adding information is not found elsewhere is the literature. Kandaswamy, Stolovitch & Thiagarajan (1976) found a main effect produced by the reviser, in that different evaluators produced revisions of different degrees of effectiveness, and also that materials revised after learner verification were an improvement. The novice revisers in Kandaswamy, Stolovitch, & Thiagarajan's (1976) study were specifically told not to use intuition. This constraint could cause problems for an experienced reviser, if it is true that such a person incorporates personal knowledge, called "tacit knowledge" or "connoisseurship" by Polanyi (1975), or if he follows Nathenson and Henderson's (1980) advice "to build up a personal 'memory store' of revisions" (p. 161). Henderson and Nathenson (1977a) showed that revisions based on

learner input, summarized by an instructional designer and used by a multi-disciplinary team who were the original course developers, produced improved learning. McAlpine (1990) and Duy (1990) both found that instructional designers reported that they based many revisions on learner data when they were provided with only learner data.

In her study of graduate students who were asked to revise each other's materials based on rules, Baker (1970) noted that results on the revised materials had improved, although the majority of changes were stylistic and not rule-based. Golas's (1982) study found that instructional designers used learner feedback and instructional editing guidelines to produce revisions of equivalent effectiveness. Burkholder (1981) described successful outcomes of using the Instructional Strategy Diagnostic Profile (ISDP) but did not describe who did the revisions. Montague, Ellis and Wulfeck (1983) described the Instructional Quality Inventory (IQI), which is similar to the ISDP, as being an effective rubric for revision, and one which was easy to teach. Dupont (1980) and Dupont and Stolovitch (1983) provided the only example of an empirical test of the revision guidelines developed by Gropper (1975). They trained non-professionals in a version of Gropper's model and found that revisions using the model did not produce increased learning of objectives, compared with versions revised using learner feedback.

McCormick (1976) found that non-professional evaluators closely involved with the designers of an Open University course produced materials which improved learning by basing their revisions on a combination of subjective student ratings and on their own intuition. Some studies have shown that revision is an idiosyncratic process, depending on the input of the reviser (Dick, 1968; Kandaswamy, Stolovitch, & Thiagarajan, 1976), while some have shown consistency among revisers with expertise in the same domain (Saroyan, 1989). Davidove and Reiser (1991) explained the lack of difference between materials revised after instructional designer input and after teacher input by the fact that the same instructional designer had implemented the revisions generated by each source of feedback. This suggests that the reviser influence was a factor, as well as the feedback sources.

The only study identified as specifically addressing the question of which types of feedback are used by experienced revisers is that conducted by Ciesla (1976). In this study, revisers were also developers of the materials. They used feedback from ten different sources and were asked to rank the sources with respect to their influence on the revisions they had produced, disregarding their own intuition as a source. Overall agreement among the seven revisers was low (41%), although the most influential source of feedback data was reported to be the qualitative feedback from teachers and the least influential was feedback from students. The instruction to disregard intuition was acknowledgement of the presence of revisers' own knowledge as an influence, but it is difficult to see how this could have been controlled. It would also have been interesting to have asked the revisers how much their own knowledge was a factor in decision-making, assuming that self-reporting is an accurate estimate of the subjects' actions.

Asking instructional designers to describe their usual behavior has not provided useful insights. Burt's (1989) survey of the usual practice of instructional designers included questions, often taken as verbatim statements from textbooks, to investigate how much they incorporated feedback data into revision. A significant number of Burt's respondents reported that they used subject-matter experts or peer reviews in addition to learners in developmentally testing prototype instructional materials and that, as a general strategy, they based revision on both the pretest and posttest data and on the comments made by test subjects. But 34% then indicated that they revised based only on obvious self-evident weaknesses; 11% claimed that they consistently followed a prescribed algorithm or heuristic; and 6% said that they revised based on intuition. Yet, paradoxically, 68% of Burt's respondents perceived that revision procedures are teachable.

The problem faced by professional revisers who are given contradictory data is identified by Lefrere (1981). Thiagarajan (1978) notes that "contradictory suggestions

from different experts (even from those in the same field) occur more frequently than one can imagine" (p 137). Further, "different developers may come up with entirely contradictory revisions (e.g., delete a paragraph vs. add more explanation to the paragraph) based on the same learner feedback" (p. 140). Thiagarajan reports using a Delphi-like technique for resolving contradictory proposals, but gives no more details.

Summarizing such a varied set of research results spread over 25 years has revealed some common threads particularly relevant to the present study.

1. Few studies examine the actions of professional revisers, preferring instead to train subjects specifically for the revision task being examined, or using the experimenter as reviser. This conclusion in Baker and Alkin's summary of research to 1973 was endorsed by Nathenson and Henderson in 1980. They criticized several studies of novice reviser behavior because the results were confounded by researcher bias and contained weak evidence that the revisions were actually controlled by try-out data.

2. Few of the studies describe how the revisions were carried out, nor how the actions of the revisers were observed. It is not surprising that textbooks contain few instructions on how to revise, when research does not identify how revisions are effected. If retrospective reporting is not to be trusted as a source of data, some other method of tracking reviser behavior must be used. Now that the think-aloud technique is an accepted data collection method, more insights into revisers' procedures are possible.

3. While several studies examined various possible sources of feedback, none of the studies investigated the case where revisers were given the opportunity to use data from both experts and learners, including test scores. Thus, instructional designers adherence to the injunctions that they are given in their training, and their selection of feedback data have not been tested.

4. More important, while several studies tried to control for reviser input, or at least note its intrusiveness, none of the studies reports that the revisers' own input was assessed or quantified in any way. Even if they are specifically told not to use their own

intuition, it is difficult to imagine that the revisers' bias or their personal knowledge could be completely excluded. These components must surely be factors in all revisions.

This present study attempts to answer some of the issues not addressed by earlier research. When professional instructional designers are studied, do they use feedback data obtained from learners and experts when they revise? What decisions do they make regarding the incorporation or rejection of various types of data? How do they deal with contradictory data or data which conflict with their own personal knowledge? Do they apply their own biases to the process of revision? Is their procedure for revision a function of the revisers' training or experience?

This last question is of interest in making the link between theory and practice and leads to the question of what procedures the revisers were instructed to follow during their training. While much of the work on formative evaluation in the 1970s seems to have focused on data collection, attempts to describe a revision heuristic have been limited. The lack of empirical evidence for an accepted revision heuristic is reflected in the instructional design textbooks produced since then. Consideration of the textbooks used in instructional design courses, which might have been used by the subjects of the present study, provides valuable clues to this connection between prescription and practice.

Textbook Prescriptions for Revision

Because Ely's (1992) survey showed that "most professional education programs preparing educational technologists do not offer a separate course in evaluation" (p.27), the educational technologists who have been trained in instructional design have probably learned a standard model of instructional design, based on the prescriptions given in one or more textbooks. If they have not followed a course specific to evaluation, their training in formative evaluation and revision must have come from these instructional design textbooks. Figure 4. Attention to revision in instructional design textbooks

egend: x indicates evidence of item.	/1	3	A Stranger	A CONTRACTOR OF			C C C C C C C C C C C C C C C C C C C	100 100 100 100 100 100 100 100 100 100	Contraction of the second		Contraction of the second seco		100 100 100 100 100 100 100 100 100 100	Carling Carling	A Less and the second s	Service in the servic	Solution of the second
1. Recognition of revision ?	X	×	x	x	x	x	x	x	x	x	x	x	x	x	x	ſ	Í
2a. Theoretical base ?						x		x			x	x	x		x	x	1
b. Need for research ?					x	x					x	x					
3. Well-articulated step ?				x	x	x	x	x	x	x	x	x		x	x		1
4a. Need for expert data ?	x	x	x	x	x	x				x	x	x	x	x	x	x]
b. Need for learner data ?	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
5. Data collection instrument ?				x	x	x	x	x	x	x		x	x	x	x	x	
6. Data handling instrument ?					x			x	x	x	x	x	x	x	x		1
7a. Revision procedure given ?					x		x	x	x	x	x	x		x	x		1
b. Reviser input ?								x	x	x	x	x		x	x		
8. Case studies or models?			x			x	x	x	x	x	x		x	x	x	x	
9. Practise ?							x		x	x	x						
10. Feedback ?							x		x	x	x						1

If revision is viewed as an ill-defined problem, it would be surprising to find a revision algorithm prescribed in instructional design textbooks, and this proves to be the case. Moreover, a review of 11 instructional design textbooks published between 1971 and 1990, shows that little instruction was offered regarding the revision process (Le Maistre, 1991, summarized in Figure 4). Questions were developed to evaluate the extent to which instructional design textbooks gave instruction on how to revise instruction. No pattern occurs if the textbooks are arranged in chronological order; over a 19-year time span, no major differences have occurred in the advice given by textbook authors to revisers. All 11 authors recognized the need for a revision step in materials development. All suggested using learners to provide data to guide revision decisions, based on the long and proven tradition of learner verification and revision in formative evaluation, but the need for expert review was identified by only 8 of the 11 authors. Of the remaining three, the books by Gropper (1975) and Popham & Baker (1971) are aimed at instructional designers and emphasize the need for feedback data from learners. The last, by Dick and Reiser (1989) is directed at practicing teachers, who would be unlikely to have either opportunity or inclination to consult experts before they revise their instruction. Only 8 of the 11 texts discussed revision as a clearly defined, well-articulated step, and this question seems to be the sieve which best discriminates among the revision suggestions given by the books examined.

Seven of the eight textbooks which gave real attention to revision suggested an instrument for collecting data, five going to the extent of proposing a method of organizing and handling the data to prepare for the reviser, and six of the texts proposing a procedure for revision. In general, this procedure followed the four-stage process described earlier as the standard model, suggesting that revisions be done after each data collection stage.

Four books told the designer/reviser to trust his own intuition in making revision decisions from the data provided, but one (Popham and Baker, 1971) strongly advised against this. In their prescription for revision, Briggs and Wager (1981) echoed Baker

(1974) in telling the reviser to assign priorities in advance. They suggested that "pretest and posttest achievement scores are usually assigned the highest priority." They gave no instruction as to how to use the priorities once they are assigned. Dick and Carey (1985) were more helpful than Briggs and Wager and advised revisers to "use the data, your experience, and sound learning principles as the bases for your revisions" (p. 232)

While seven of the eleven books supplied case histories, only four of them allowed for any practice and feedback. Finally, only three of the revision prescriptions were based on a theory which is also described in the book, and only three texts suggested a need for research on revision.

In 1991, Gustafson published a survey of instructional development models which covers the 10 years after Andrews and Goodson's (1980) survey. Ten of the 12 models described in this survey refer to revision as a distinct step in formative evaluation. But an update to the review of textbooks, based on five more texts published since 1990 and shown in the last five columns of Figure 4, does not change appreciably the picture of textbook instruction to students of instructional design. It should be noted that these later books could only form part of the training of recently graduated instructional designers, and that these recent graduates do not form part of the population under consideration in this study.

It appears that instructional design textbooks, which are otherwise exemplary, are weak in the area of revision. Even if they recognize that revision is a discrete and important step in systematic materials development, authors are likely to give less instruction, practice and feedback on this step in the process than in other areas. Prescriptions exist as to how to collect data from learners on the effectiveness of teaching materials, and think-aloud procedures have proved useful in collecting feedback from experts (Saroyan, 1989). Designers are told that these data should be used to guide revisions, but there are few suggestions as to how this should be done, with the exception of texts such as Gropper (1975) and Dick and Carey (1991). Yet instructional designers are charged with revising instruction on a daily basis. The Association for Educational Communication and Technology lists "Generate specifications for revision based on evaluation feedback" (p. 15) as one of its core competencies for a professional designer (Beery et al., 1981).

This analysis of textbook instruction in revision was conducted to find out how much help is given by textbooks to neophyte revisers, and to provide an estimate of the training received by the instructional designers who might be involved in the present study. If their training may be summarized as the brief injunction to collect feedback data and use this information to revise, what procedure do they actually follow?

Summary

The literature in this chapter was surveyed under four main themes. First, revision was identified as an ill-defined problem, with similarities to the act of revision in composition. Placing revision within the framework of human problem solving provided a theoretical framework for the present study.

Second, establishment of a well founded theory base has meant that the methodologies resulting from the theory could be applied to the research questions for this study. The think-aloud technique was described and proposed as a useful strategy for revealing the cognitive activity of problem solvers, notably those who were experienced in the domain of problem under investigation.

Third, the instructional design literature was examined for research on the actions of revisers. This revealed the need for research designed to observe the actions of experienced instructional designers while they were in the act of revising. Such research would be related to the training of instructional designers, when they are told to use feedback from experts and learners, and should also be designed to examine the impact of the revisers' personal knowledge on their actions.

Fourth, to establish a profile of the training that might have been followed by the community of experienced instructional designers, a selection of popular instructional

design textbooks was examined for the information they gave regarding revision. It was found that, with some exceptions, instructional design textbooks generally give little information on how to revise. This might be explained by the lack of research already identified in the third section of the review.

The next chapter will describe the design for the study and explain the methodology used.

CHAPTER 3: DESIGN AND METHOD

This study sought to investigate the actions of professional instructional designers when they were asked to revise a module of self-instructional text. In one condition, with feedback, revisers were presented with the results of evaluation from a variety of sources recommended in textbooks: learner feedback (oral); learner data (test scores); expert feedback (oral data), and were asked to revise the modules so as to produce an improved version of the text. Revisers were not specifically told to use these data, nor forbidden to apply their own knowledge base, so that all these sources could contribute to the revision. In effect, subjects in this condition were asked to perform a "meta-evaluation" (Kandaswamy, undated, cited in Nathenson & Henderson, 1980) of the suggestions they were given, qualified by their own knowledge of instructional design and by their experience.

In the second condition, without feedback, subjects were applying only their instructional design knowledge and their previous experience in revision, presumably based on the training they had received. Since no review data were available to them, they were forced to act as both reviewers and revisers in the without feedback condition.

Design

The independent variable investigated in this study was the feedback condition, which had two levels, with feedback and without feedback. The dependent variable was the number of revision statements linked to each source of feedback data. The study used a repeated measures design, with each subject given both levels of the feedback condition (F1, with feedback, and F2, without feedback). To enable investigation of any confounding of order effects with treatment effects, the design was counterbalanced with respect to order of feedback (feedback first, at time T1, or feedback second, at time T2), and subjects were nested within this order variable. This was necessary because no

research exists to predict whether the provision of feedback during one revision will produce a learning effect for any subsequent revision carried out in close time proximity but without feedback, nor whether a data-free revision task affects subsequent choices among data sources in later revision tasks.

Two modules were used $(M_1 \text{ and } M_2)$, so that the feedback variable could be investigated independently of the content of the modules being revised. The modules were not randomly selected, for the reasons outlined below under "Stimulus materials", but the design allowed any effects produced by any differences in the modules to be identified. It also controlled for individual preferences among subjects for data sources, or for differences in inclination to use tacit knowledge. Each subject was self-matched across treatments, thus reducing the minimum number of subjects needed. In view of this economy, it was possible to assign two subjects to each of the four levels of module and feedback (S₁ through S₈).

In parenthesis notation, this design may be expressed as:

$$S_2(O_{F_T} \times O_{M_T}) \times F_2$$

where S, the subject, is nested with the order of administration (O) of two factors: OF, the order of the feedback condition (with or without feedback), crossed with OM, the order of administering the module (M1 or M2), and T is the time of administration. This is all crossed, in turn, with the two levels of the feedback condition, F. This design is summarized in Table 1.

Table 1.

Summary of Research Plan

Condition	Subject	Time 1 the	en Time 2
1	(S ₁ , S ₂)	Module 1 With feedback	Module 2 No feedback
2	(S3, S4)	Module 2 No feedback	Module 1 With feedback
3	(\$5, \$6)	Module 2 With feedback	Module 1 No feedback
4	(S7, S8)	Module 1 No feedback	Module 2 With feedback
	-		

Method

Data collection for this study consisted of two phases. Phase 1 was the review phase, involving the collection of feedback data from experts and learners, and phase two was the revision phase.

In phase 1, two modules were selected for revision, according to criteria which will be outlined. The modules were reviewed by experts and learners to provide feedback data that could stimulate revision. Although the review phase of formative evaluation was not the focus of the study, it was an essential first step before phase 2 could be started.

In phase 2, the design described above was implemented. Eight practicing instructional designers $(S_1 - S_8)$ were randomly assigned to one of the four conditions outlined in Table 1, and asked to revise the modules while thinking aloud. Their verbalizations were tape-recorded by the researcher and a coresearcher, and the audiotapes were transcribed and segmented. A coding scheme was developed to identify the revisions and the source which had prompted the revisions, and the segmented protocols were coded. The method is represented graphically in Figure 5. In addition to providing the

.

 $\widehat{}$

think-aloud data, each subject was told to write anything they wanted, and each was debriefed during a retrospective interview.

Phase 1: Collection of Feedback Data

The design of this study involved the preparation of four different packages of materials depending on the condition and time of administration. The first task was to choose two equivalent modules of instructional material as stimulus material, and the second was to obtain feedback data from experts and learners on each module.

Selection of Stimulus Material

The three areas most directly involved in formative evaluation are the teaching profession, the publishing industry and the training departments of large companies. Each area develops materials which lend themselves to research in varying degrees. In teaching, materials developed by teachers are constantly undergoing revision while they are in use. Revision in stand-up teaching, more precisely called non-replicable instruction, introduces confounding variables such as the performance and style of the teacher, and does not readily lend itself to research.

The present study avoided the problem of non-replicability by using written materials. Replicable instruction, such as that found in publishing, training, or university course materials is susceptible to research projects in a naturalistic setting. In publishing, revisions are conducted on print material by the original authors or by professional editors who are not usually trained in systematic materials development. In business and industry, one of two cases usually applies. The training material may be systematically developed by instructional designers who also revise the materials after testing and before they are used on a large scale. Otherwise, the training material may be developed by a content specialist in consultation with an instructional designer and revised by the instructional designer before general use. In a university context, materials are frequently developed by professors (content specialists) for a particular course. If they are subsequently revised, the revision is performed by the original author. This last case, the development of materials for a university course, was examined in this study, except that the authors gave permission for the modules to be revised by instructional designers, so that the behavior of professional revisers could be examined.

The stimulus materials comprised two units of self-paced instructional material, part of a collection of 28 such modules, totalling 153 pages, written by subject-matter experts in chemistry as part of an introductory chemistry course for first year undergraduate students in Arts or Science. The first module, M₁, was called "The diet-cancer relationship" (Fenster, Harpp, & Schwarcz, 1990b) and the second, M₂, was called "Artificial sweeteners" (Fenster, Harpp, & Schwarcz, 1990a).

The modules were not contrived for experimental purposes and did not constitute "an intentionally sloppy prototype" (Thiagarajan, 1978, p.141). But although they had already been used with several cohorts of learners, they had never been formatively evaluated, and had not been professionally published. There appeared to be sufficient flaws in format, layout and typography to consider them as draft materials.

Choosing university level materials provided a pool of learners to give feedback data and access to experts familiar with the target population. The content of the modules made it unlikely that revisers would have expertise in the subject matter of the materials beyond that of any educated lay person, so that they would not bring subject-matter expertise to the revision task, reflecting the situation common in their usual practice.

The modules were selected for their length (about 30 - 60 minutes of instruction) and so as to be as alike as possible. The size of the modules was long enough to constitute a valid learning task, yet short enough for experts and learners to review without fatigue. Equally important, they were short enough for revisers to be asked to make revisions in one session. Both modules were written by the same subject-matter experts, for the same audience, and neither module acknowledged material adapted from other sources. The two modules were as closely matched as possible. One module was 5 pages and the other 6 pages long. Both scored at Grade 14 on the Flesch-Kincaid test, with Reading Ease Scores

of 37 and 38 respectively on the Flesch scale. Both were closely associated with diet, rather than "pure" chemistry, so that it was appropriate to use dietitians as subject-matter experts to review them. The learning task was essentially the same for both modules, since both involved complex chemical names, and neither was based on case histories, as were some other modules in the package. Finally, there were similarities in surface characteristics; both included a chart, typographical errors, one illustration each, and the same format. Counterbalancing the conditions under which the stimulus materials were administered was planned to allow for any residual differences between the two modules.

That the materials have been developed by subject-matter experts represents a common case in practice. The fact that they were not developed according to a systematic design model is irrelevant to this study, whose focus is the actions of instructional designers and the choices they make among data sources when they revise. As well, the content of the materials is irrelevant, beyond the fact that it was unfamiliar content to the subjects used as revisers in the study; it was the kind of feedback data (learner feedback, test scores, and expert feedback) and the decisions prompted by these data that were of interest.

Obtaining Feedback Data on Stimulus Material

Once the two sets of stimulus materials had been selected, it was necessary to have each of them reviewed by learners and experts to provide revisers with feedback data. As part of a programmatic series of studies in formative evaluation, feedback data had already been collected on one module, "The diet-cancer relationship" (Rahilly, 1991; Israeloff, 1992, Tremblay, in progress). The researcher collected data on the second module, "Artificial sweeteners."

The collection of feedback data by different researchers may be seen as a limitation to the validity of the data, but the concern of this study was the revisers' selection among data sources. In addition, care was taken to use matched procedures for collecting feedback data on each module, to the extent of using the same script when talking to

subjects at all stages of data collection for each module. These procedures will now be outlined, and any differences between them will be noted. Further details of data collection on "The diet-cancer relationship" may be found in the relevant theses.

Learner feedback: test data. Learners to provide feedback for both modules were solicited so as to be as much like the target audience for the modules as possible; that is, undergraduate students in either Arts or Science who were not majoring in Chemistry, and who had not already taken the course for which the modules had been written.

A pretest and posttest were developed for each module. Since there were no objectives provided with the modules, it was necessary to identify the key concepts to be tested. A chemistry teacher, an instructional designer and a researcher read through each module and highlighted the major concepts. Concepts identified by at least two of the three were retained, and a researcher developed test items based on these concepts. Items were pilot tested by five fellow researchers, and, in the case of "The diet-cancer relationship", also pilot tested by a representative group of learners. The tests developed for "The dietcancer relationship" consisted of 9 true/false pretest questions, matched with 9 alternate form true/false posttest questions, with an additional 14 multiple choice posttest questions, for a total of 23 questions. The test data for "Artificial sweeteners" consisted of a pretest of 8 true/false questions, and a posttest of 8 alternate form true/false questions and 15 multiple choice questions, for a total of 23 in the posttest. The format and presentation to learners of both teaching material and tests were identical for both modules. For "The diet-cancer relationship", 17 of 31 available sets of test scores were chosen at random. Seventeen learners wrote the pretest for "Artificial sweeteners", studied the materials and wrote the posttest. The percentage of learners who were successful was calculated for each pretest and posttest item in each module. Results for the pretests and posttests are given in Table 2.

Table 2.

 \bigcirc

 \bigcirc

Percentage of Learners Who Answered Items Corectly for the Pretest and Posttest for Each Module

Item #		dule 1 : 17)	Module 2 $(n = 17)$		
	Pretest	Posttest	Pretest	Posttest	
1	26	88	59	100	
2	85	94	100	41	
3	97	85	59	94	
4	79	94	41	94	
5	65	53	59	65	
6	44	79	71	65	
7	74	79	29	76	
8	47	47	71	76	
9	38	85		76	
10		82		47	
11		79		71	
12		68		76	
13		44		41	
14		94		35	
15		79		59	
16		91		76	
17		91		71	
18		71		71	
19		35		88	
20		88		59	
21		76		82	
22		88		76	
23		100		88	

Learner feedback: oral data For both modules, oral comments were collected from learners under three conditions: a single learner (n=1); a pair of learners working together (n=2); a small group of learners working together (n=5). After writing the pretest, each group was instructed to learn the materials as well as they could and to make any comments, positive or negative, regarding the modules as they went through, so that their comments could be audiotaped. After they had finished the posttest, these learners were debriefed according to an agenda developed from their comments while they were working.

For "The diet-cancer relationship", oral comments were transcribed verbatim, segmented, and coded (see Rahilly, 1991). For "Artificial sweeteners", the researcher listened to the audiotapes and extracted comments of the sort coded by Rahilly as problems identified; evaluation statements, both positive and negative; revision suggestions, since these statements provide the most information to the reviser. The script used by the administrators for greeting learners, giving task instructions, clarifications and thanks was identical for each module (see Rahilly, 1991, Appendix B, for these scripts). Strict attention paid to these procedures helped to ensure that data sets were equivalent for the two modules. Comparison of the feedback data showed that there were comparable numbers of statements coded as problems, evaluations and revisions for each module, and provided further evidence for the fact that data collection methods were parallel, and that the two modules were similar.

Expert feedback: oral data. The feedback data provided by expert reviewers were the subject of two theses in the program of research on formative evaluation (Israeloff, 1992, Tremblay, in progress), so that feedback data existed from experts for the module "The diet-cancer relationship." Six experts provided review data for "The diet-cancer relationship." The three subject-matter experts (SMEs) were practicing therapeutic dietitians, each of whom had post-baccalaureate certification, belonged to the professional corporation and had more than eight years related work experience. The three audience experts (also called target population experts, or TPEs) were university teachers who had at

least five years teaching experience with the type of learner for whom the materials were intended. All reviewers had previously reviewed instructional materials.

Reviewers were instructed to read through the materials aloud and to make any comments regarding the content of the materials and its appropriateness for the intended learners. They were told to say aloud any thoughts that occurred to them and any comments that they wrote down. They were prompted by the administrator if they stopped thinking aloud. The think-aloud protocols for "The diet-cancer relationship" were transcribed and segmented and coded in the same way as the learner feedback had been processed. Ambiguities during the think-aloud were clarified during a debriefing session.

To collect feedback on "Artificial sweeteners", this data collection process was replicated by the researcher. One subject-matter expert and one target population expert whose qualifications matched those of the six reviewers already described were used. The same procedure was followed and the same script used in communicating with the reviewers. The researcher identified comments made by these two experts which fell into the categories identified by Israeloff and Tremblay as problems identified, evaluation statements, both positive and negative, and revision suggestions, because of their usefulness to revisers. The frequency of segments in each of these codes was graphed for the 6 reviewers for "The diet-cancer relationship" and for the two reviewers of "Artificial Sweeteners" The graphs were compared in order to choose one subject-matter expert and one target population expert from "The diet-cancer relationship" whose pattern of data resembled most closely the data from the expert reviewers for "Artificial sweeteners" (see Figure 6).

This figure shows the problems identified (PI), problem elaborations (PE), problem reiterations (PR), evaluation statements (ES) and revision suggestions (RSG) made by the target population expert and the subject-matter expert selected from Module 1 (Module 1 TPE and Module 1 SME) and the same segments for the reviewers for Module 2 (Module 2

TPE and Module 2 SME). The small number of revision suggestions made by the subjectmatter expert for Module 1 should be noted.

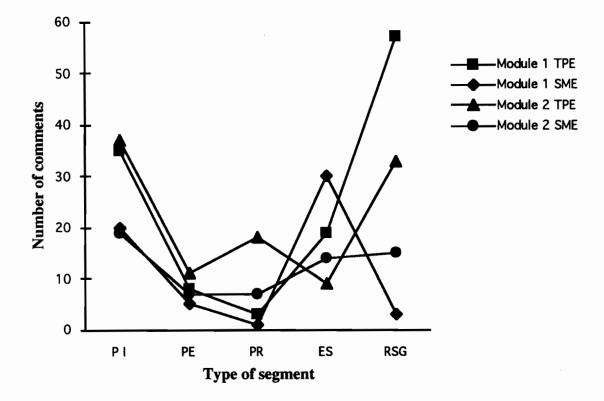


Figure 6. Comparison of reviewers for Module 1 and Module 2

In summary, this procedure was used to produce parallel data sets for the two modules under investigation, and contained as many safeguards as possible to ensure the equivalence of the feedback data for each module. Learners, subject-matter experts and target population experts provided information appropriate to their abilities, specifically, for each module:

1. Learner feedback (oral). Oral comments from concurrent verbalization during the learners' reading of the text, and oral comments made during the debriefing;

2. Learner feedback (test scores). Pretest and posttest scores on a criterion test constructed for the study and composed of criterion items linked to major concepts in the module;

3. Expert feedback. Comments on content and appropriateness made by a subjectmatter expert (abbreviated to SME) and an audience expert (also called a target population expert, or TPE) who were audio-taped while they were thinking aloud, and during a debriefing session.

Further detail on the expert review studies may be found in the relevant theses.

The data produced by each of these reviews from learners and experts were synthesized by the researcher, and prepared for the revisers' use. This will be described in detail under "Materials for gathering revision data." It should be noted that the input of instructional designers was deliberately omitted from this review phase so that such data would not contaminate the input of instructional designers during the revision stage.

Assembling the packages of materials was the last stage in phase 1 of the data gathering. In phase 2, the modules were revised, and this stage will now be described.

Phase 2: Revision of Modules

Subjects Used as Revisers

The revisers were eight instructional design practitioners, whose training had involved instruction in formative evaluation, and whose jobs involved revision of print materials. Instructional designers, rather than writers, or the original authors, were used as revisers, since instructional designers are trained in the process of using feedback data for revision, and since there is evidence that neither professional writers (Graves & Slater, 1991), nor the materials' original developers (Nathenson & Henderson, 1980) are the best revisers. The think-aloud method proposed for collecting data is time-consuming, and produces a large and detailed body of oral data for analysis. For these reasons, the number of revisers was limited to eight, although the design maximized the effect of a small number of subjects. The criteria for selection of the subjects (i.e., trained and experienced instructional designers) allowed conclusions to be drawn for this population. Solicitation of subjects. A list of 14 potential subjects was generated by peer nomination from other experienced instructional designers. All subjects met the following criteria:

1. Subjects had academic training in instructional design, and each had a degree in instructional design. While different university programs emphasize different aspects of the instructional design process, the attainment of a degree in the area suggests a common level of expertise and a commitment to basic principles of the systematic design process, specifically for this study, the prescription to use feedback data to make appropriate revisions.

2. Subjects had practiced instructional design for at least 5 years. While many of the existing revision studies have attempted to train novices in revision, this study was designed to examine the actual practice of experienced revisers. Expertise in revision is difficult to define, since there is no specific training for revisers beyond the regular training of an instructional designer. But, based on the content of standard textbooks (Le Maistre, 1991), this training includes the injunction to revise materials based on formative evaluation data. It was assumed that since expertise improves with practice, five years of experience would probably assure an appropriate level of competence.

3. Subjects were all actively engaged in development projects that included formative evaluation and revision. Since all subjects were either employed by large companies or self-employed as consultants, this criterion was also met.

Potential subjects were telephoned and asked for their help in a research project investigating the revision of print materials under different conditions. They were told that they would be required to verbalize their thoughts while revising, that there would be two sessions of between two and three hours each, and they were offered an honorarium for their participation. The first eight instructional designers reached by telephone agreed to participate, and confirmed that they met the minimum criteria for participation. A retrospective interview after the second session allowed the researcher to explore specific

features of each reviser's training. Profiles of the subjects will be described in Chapter 4 as part of the results of the study.

Procedure

Outline of procedure. The eight subjects were randomly assigned to one of four conditions, the assignment being predicated by the order in which the subjects were reached by telephone. When they arrived, subjects were greeted by two administrators (the researcher and a coresearcher) and given a folder of materials whose contents depended on the condition under investigation. Formative evaluation is traditionally considered to be an iterative process, and models of instructional design advocate collecting feedback and using it for revision at each stage of materials development. While ideally, revisers should be given data from different sources at various stages of the development of materials, so that revisions may be made sequentially (e.g., Baker, 1974), practical constraints mean that this is rarely followed in practice (Dick & Carey, 1990; Dick & Carey, 1991). In this study, revisers were provided with all feedback data at once during the "with feedback" condition, partly because this is often the case in reality, and partly to force the reviser to be selective and to prioritize, since this is a focus of the research questions.

The researcher explained the contents of the folders, and subjects were instructed to think aloud while they worked through materials in the order presented in Table 3. Juice and snacks were available to subjects, who were told that they could take a break in the think-aloud whenever they chose. If the subjects seemed to be tiring, losing concentration, or becoming overwhelmed by the task, the researcher suggested that it was time for a break. During these breaks, the administrators took care that conversation did not allow the subject to lose the thread of the task, but that the subject did not make any comments specific to the task either. After the think-aloud, subjects were debriefed and asked for demographic information. Materials for gathering revision data. Packages of materials were assembled for the revisers and for the administrators. The contents of these packages varied depending on the condition employed, but were consistent for each subject within each condition. The researcher used prepared scripts for each session, and the only variations in the scripts were caused by the different treatment levels of feedback and time of administration. Because of the different levels of treatment involved, this produced several different versions of each script, resulting in a large package of materials. The package of materials is summarized in Table 3, where:

A refers to the contents of the administrators' packages,

S refers to the contents of the subject's package,

M1 is module 1: "The diet-cancer relationship",

M2 is module 2: "Artificial Sweeteners",

T1 and T2 refer to the first and second sessions,

w means the with-data condition, and

n the no-data condition.

Appendix B contains copies of the materials in the different packages, labelled according to the code in Table 3. For conciseness, several components have been compressed on to one page. Not all versions of the draft materials are included. This will be noted where appropriate.

The experimental procedure will now be described, using each component of the packages and its function as an outline.

<u>1 Script for icebreaker (A only)</u>. At the first session, subjects were thanked for their participation, told that the study was part of an ongoing research project on formative evaluation and that this piece of the project would investigate how instructional designers revised under different conditions. They were reminded that they had agreed to revise one module in this session and one in the next [component 1 (T1)].

Table 3

0

Summary of Materials Packages for Each Condition, Labelled as in Appendix B

To:	with	T1 (w) feedback no	T1 (n) feedback	T2 (w) with feedbacl	T2 (n) k no feedback
A	Script for icebreaker	1 (T 1)	1 (T1)	1 (T2) w	1 (T2)n
A,S	Summary of activities	2 (T1)	2 (T1)	2 (T2) w	2 (T2)n
S	Consent form	3 (T1)	3 (T1)		
A,S	Task description	4 (T1)w	4 (T1) n	4 (T2) w	4 (T2)n
Α	Check for understanding	5	5	5	5
A,S	What is the think-aloud?	6	6	6	6
Α	Warm-up instructions	7 (T1)	7 (T1)	7 (T2)	7 (T2)
S	Think-aloud practice	8 (T1)	8 (T1)	8 (T2)	8 (T2)
S	Working copy of module	9 (M1/ M2)	9 (M1/M2)	9 (M1/M2)	9 (M1/M2)
S	Original draft of module	10 (M1/M2)	10 (M1/M2)	10 (M1/M2)	10 (M1/M2)
S	Draft + expert feedback	11 (M1/M2)		11 (M1/M2)	
S	Draft + learner feedback	12 (M1/M2)		12 (M1/M2)	
S	SME qualifications	13 (M1 / M2)		13 (M1 / M2)	
S	TPE qualifications	14 (M1 / M2)		14 (M1 / M2)	
S	Learner characteristics	15 (M1 / M2)		15 (M1 / M2)	
S	Pre/posttests	16 (M1/M2)		16 (M1 / M2)	
Α	Prompting words	17	17	17	17
Α	Tracking sheet	18 (M1/M2)	18 (M1/M2)	18 (M1/M2)	18 (M1/M2)
Α	Debriefing script	19 (T1) w	19 (T1) n	19 (T2) w	19 (T2) n
Α	Goodbyes	20 (T1)	20 (T1)	20 (T2)	20 (T2)
S	Honorarium receipt			21 (T2)	21 (T2)

At the second session, subjects were thanked for returning, reminded about the research project and told that the conditions were slightly different; that they would be given feedback data [1 (T2) w], or that there would be no feedback [1 (T2) n], depending on which level of feedback had been involved in the first session.

<u>2</u> Summary of activities (A and S). At the first session, all subjects were asked to read a short summary of the activities [2 (T1)]. During the second session, the summary varied from the first session only if there was feedback [2 (T2) w], or if there was no feedback [2 (T2) n] during the second session.

<u>3 Consent form (S).</u> Subjects were asked to sign the consent form [3], if they agreed to what it said.

<u>4 Task description (A and S).</u> There were four versions of the task description [4 (T1) w, 4 (T1) n, 4 (T2) w, and 4 (T2) n]. Each one contained the same first paragraph, namely:

This unit has been developed for use in a first-year chemistry course for both Arts and Science students who are not majoring in chemistry. The original draft was written by a subject-matter expert who is also a teacher. This person is dissatisfied with the effect of the materials on students. You are being brought in as a consultant to revise the unit. Please make **any** revisions that seem necessary to you.

The "with feedback" condition then explained that learner data were provided on green paper, and expert data on yellow paper. Each version also contained a brief introduction to the think-aloud procedure. Subjects were asked to read this silently and to ask for any points of clarification.

<u>5 Check for understanding (A).</u> When subjects affirmed that they had read the task description to their satisfaction, they were asked: "Now that you've read the instructions, would you please tell me your understanding of the task?" [5]. It will be noted that "revision" was not defined for the subject. Since it is known that problem-

solvers solve the task they represent for themselves (Flower & Hayes, 1980, p. 22), it seemed important to have revisers express the problem for themselves, while still allowing the researcher to clear up any gross misunderstandings of the task, before allowing the subject to proceed. This procedure, combined with a similar question during the retrospective interview, also allowed the researcher to check the stability of the task definition across the duration of the task. Asking the same two questions at each session allowed the same stability check across the two time conditions.

<u>6 What is the "think-aloud"? (A and S).</u> At each session, subjects were asked to read a page of information on the think-aloud process [6] and allowed to ask any questions.

7 Warm-up instructions (A). Subjects were given oral instructions [7 (T1) and 7 (T2)] about how they were expected to think aloud.

<u>8 Think-aloud practice (S).</u> It has been shown that subjects perform better during a think-aloud task when they are given the opportunity to practice beforehand (Ericsson & Simon, 1980; Flaherty, 1984). The warm-up exercise should allow subjects to practice listening to their own voices and should comprise a meaningful task, and various tasks have been used in this program of research (Duy, 1990; Israeloff, 1991; Saroyan, 1989; Tremblay, in progress). For this study, it was decided to ask subjects to perform a revision task on text similar to the actual stimulus materials. Two modules ("Monosodium glutamate" and "Preservatives") were selected from the same package of materials as the modules that subjects would subsequently revise, and the first page of each of them was reproduced in the same format as the subjects would see during the actual task.

Subjects were told that on the actual task, they would take as much time as they needed, but the practice would be restricted to 3 minutes. This allowed the researcher to estimate the subject's level of comfort with the process and to allow extra time, if necessary, or to stop the practice sooner. During the practice, the researcher was also able to coach subjects who were uncomfortable with the task.

During the first session, subjects practiced thinking aloud on "Monosodium glutamate" [8 (T1)], and were offered "Preservatives" [8 (T2)] for the second session, if they felt they needed a refresher in the technique. None of them chose to use it, and the researcher did not press the offer, as the two sessions were close enough for each subject that it was unlikely that they had forgotten the technique.

<u>9 Working copy (S).</u> Subjects were given a copy of the module [9 (M1) or 9 (M2)] printed in the centre of legal size paper, oriented in landscape layout to allow them space around the text for making any annotations that they chose. (This does not appear in Appendix B).

<u>10 Original draft of module (S)</u>. Subjects were also given a copy of the original module [10 (M1) or 10 (M2)] so that they could see the actual materials given to students and could make any comments on layout and presentation that would not be apparent from the working copy.

11 Draft with expert feedback (S). The text of each module was printed in the centre of legal size paper, oriented in landscape layout, and the identified comments (i.e., evaluation statements, problems, possible revisions or statements of knowledge) were printed on either side of the text. Subject-matter expert comments were printed in the left column, and target population expert comments in the right column. The module had no explicit objectives to which comments could be referred (as recommended in Dick & Carey, 1991), so each comment was located alongside its referrent line. When reviewers referred to specific words of the text, these words were identified with an asterisk. General comments from the expert reviewers were inserted at the end of the module . The only editing of comments was done for the sake of brevity. These data were reproduced on yellow paper [11 (T1) w, and 11 (T2) w]. (N.B.: These and all other materials originally presented to subjects on legal size paper have been reduced to quarto size for inclusion in Appendix B).

<u>12 Draft with learner feedback (S)</u>. The text of each module was printed in the middle of legal size paper, in landscape layout. Pretest and posttest success rates were recorded as percentages for each question and were printed in two columns to the left of the text and alongside the relevant passage being tested. Learner comments were aggregated and printed to the right of the text alongside the relevant text. Specific words being addressed in the comments were identified by an asterisk and where more than one learner had made the same comment, the number was noted in parentheses. General comments from the learners were inserted at the end of the module. This draft was reproduced on green paper.

The choice of colors for feedback data was arbitrary, but the use of color was deliberate, to make it easier for subjects to organize the large quantity of paper involved in the "with data" condition [12 (T1) w, and 12 (T2) w].

<u>13 SME qualifications (S)</u>. The qualifications and data collection method for the subject-matter expert (SME) were copied on yellow paper and provided for revisers to consult [13 (T1) w, and 13 (T2) w]

<u>14 TPE qualifications (S)</u>. The qualifications and data collection method for the target population expert (TPE) were copied on yellow paper and provided for revisers to consult [14 (T1) w, and 14 (T2) w].

<u>15 Learner characteristics (S).</u> On green paper, revisers were given information on the learners, and on the methods of collecting test data and oral data [15 (T1) w, and 15 (T2) w].

<u>16 Pre/posttests (S)</u>. Each test was copied and provided for revisers [16 (T1) w, and 16 (T2) w].

<u>17 Prompting words (A).</u> A list of prompting phrases [17] was prepared for the researcher's use to encourage verbal production from the subject. If the reviser stopped talking for more than 10 seconds, the researcher prompted with the first phrase, and could go down the list using successively stronger prompts if silences continued.

<u>18 Tracking sheet (A).</u> Both researchers had a tracking sheet, modified from Duy (1990), and based on the working copy [9] given to the revisers. Using the same format as the revisers' working copy made it easier to record the location of the revisers' written comments, highlightings or underlinings, and page turnings. In the "with data" conditions, the tracking sheets had four columns drawn on the right hand side so that the researchers could track reviser references to test data, learner comments, SME comments and TPE comments [18 w]. In the "without data" condition, the tracking sheet [18 n] was identical with the revisers' working copy. The tracking sheet provided part of the agenda for the retrospective interview. (This does not appear in Appendix B).

<u>19 Debriefing script (A).</u> There were four versions of the debriefing script for the retrospective interview, depending on the levels of feedback and time [19 (T1) w, 19 (T1) n, 19 (T2) w, 19 (T2) n]. Each debriefing session began with a period of clarification and elaboration of points made by the reviser during the think-aloud, prompted by the notes taken on the tracking sheets by both administrators. This compensated for the absence of probing during the think-aloud. Each version then included questions to focus the reviser away from the task just completed and towards his usual practice, and then to compare the two tasks, making reference to the individual data sources in the "with data" condition [19 (T1) w and 19 (T2) w]. This would allow a later comparison of each reviser's actual practice with his perception of the procedure he usually followed, since there is evidence (Duy, 1990) that revisers do not accurately estimate the attention they pay to each data source.

All debriefing sessions asked revisers to report any problems they had experienced and to rate their familiarity with the materials and the audience. To find out if their task definition had changed during the revision, they were asked to answer the question "If you go back to the office and have to describe this task to a colleague, how would you describe the task you've just completed?" They were given the opportunity to give general opinions and to ask any questions about the research. In the first session, if these questions

compromised the research design, they were noted and deferred until after the second session. Debriefing during the second session ended with questions relating to the revisers' training, experience, commonly used reference materials, and professional affiliations [19 (T2)]. Answers to these questions supplemented the information obtained during the telephone solicitation of subjects and verified that the criteria for participation had been met.

<u>20 Goodbyes (A).</u> Revisers were thanked for their participation and either the second appointment was confirmed [20 (T1)], or they were given an honorarium [20 (T2)].

<u>21 Honorarium receipt (S).</u> After the second session, revisers were asked to sign a receipt to acknowledge receipt of a cheque for the honorarium [21].

To summarize the contents of each package of materials:

The subject's package comprised: Summary of activities, consent form (T1 only); Task description; What is the "think-aloud"?; Think-aloud practice; Working copy of module; Original draft of module; Draft with expert feedback (with data condition only); Draft with learner feedback (with data condition only); SME qualifications (with data condition only); TPE qualifications (with data condition only); Learner characteristics (with data condition only); Pretest and posttest (with data condition only); Honorarium receipt (T2 only).

The administrators' package comprised: Script for icebreaker; Summary of activities; Task description; Check for understanding; What is the "think-aloud"? Warm-up instructions; Prompting words; Tracking sheet; Debriefing script; Goodbyes.

Packages of materials were prepared ahead of time and placed in different folders for each session. Care was taken to organize the materials as carefully as possible, partly because of the complexity caused by the different instructions for different conditions, and partly to make it easier for the subjects to understand the large amount of material in the "with data" condition.

Equipment. Each session was audiotaped with two tape recorders using professional quality PZM microphones. The times for turning or changing tapes were staggered so that subjects were not required to pause during the think-aloud, and comments missed on one tape were available on the other. During tape changing, one researcher continued the tracking while the other changed tapes.

Role of researcher: during the revision task. The researcher sat at a 90° angle to the reviser so that the reviser's actions could be noted as unobtrusively as possible. The researcher audio-taped comments made by revisers during the think-aloud to provide verbal protocols for later analysis, and also tracked the revisions on the tracking sheets by marking when data sources were being consulted and noting any unclear comments made by the reviser. The location of these unclear comments was marked and formed the agenda for the first part of the retrospective interview after completion of the revision task. These uses of the tracking sheet provided later verification of the information provided by the think-aloud process. The researcher verified at the outset that subjects had understood task instructions, and answered questions regarding the task. If pauses in speech production were longer than ten seconds, the researcher reminded revisers to verbalize their thoughts continuously. Otherwise, subjects were not prompted, and researcher comments were verbal or non-verbal expressions of encouragement and support, such as "Uh-huh", "OK", "I see", or nodding of the head.

Role of researcher: after the revision task. After the think-aloud, the researcher conducted a retrospective interview to debrief the subjects to clarify any of the comments made during the think-aloud, to ask subjects their usual revision practice, and to give the subjects an opportunity for more general comments. The agenda for the first part of the debriefing was developed from notes taken on the tracking sheets during the think-aloud process, and was designed to fill omissions and clarify ambiguities. In the second part of the debriefing, the researcher followed a script of questions to ascertain the subjects'

usual practice and perception of the task just completed. In addition to this, the second session for each subject ended with an examination of the subject's background and credentials, and gave subjects the opportunity to ask any questions relevant to the study.

Role of coresearcher. Another member of the research team acted as coresearcher and assisted in data collection. The coresearcher sat across a large table from the reviser to monitor the tape recorders and microphones and to change tapes. He also tracked the revisions, following the same procedure as the researcher, and comments noted by the coresearcher formed part of the agenda for the retrospective interview.

Data Analysis

Transcription. The think-aloud data from the eight subjects were transcribed verbatim by qualified transcribers, and each transcript was checked for accuracy against its audiotapes by the researcher. Verbatim reading of the text was enclosed in quotation marks (""). Hesitations were marked with two dashes (--). Pauses corresponding with subjects' speech bursts were punctuated with a comma (,), and longer pauses were marked with an ellipsis (...). Emphasis was added with boldface type. When the subject wrote anything, the comment was underlined, and when it was impossible to understand what was being said, asterisks were used (***). After the transcript had been corrected by the researcher, there were few such gaps, representing, at most, a one- or two-second break in the text, often associated with some movement close to the microphone.

Segmenting. The transcripts were then segmented into "meaning units" (Rahilly, 1991), using the following definition of a segment:

A segment is a meaningful unit which may include a tensed (conjugated) verb, a participle, an infinitive, or an understood (implied) verb and any bound adjuncts. Bound adjuncts are phrases (subject and complete verb absent) or clauses (subject and predicate present) beginning with a preposition or a subordinate conjunction

which modify or add meaning to the unit or segment (Rahilly, Weston, McAlpine, 1991, p.2).

This definition has proved useful in producing segments containing one main idea, which may subsequently be coded, yet it is not so restrictive that ungrammatically expressed (but theoretically important) ideas are lost. It was found to be helpful to have a rationale for segmenting established *a priori*, and made it easier to accommodate the various speech idiosyncracies of the subjects. Segments were marked with a double slash (//), each segment was typed on a new line, and each segment was numbered.

Stability of the segmenting was checked by resegmenting the protocols after a gap of at least three months. In typical transcripts of 800 - 1200 segments, no more than 20 segments were ever changed. Most commonly, these were changes where the original segmenting had been too fine grained, and two fragments needed to be combined to satisfy the definition of a segment.

Coding.

Development of coding scheme

Coding schemes reflecting the problem-solving nature of revision (Hayes, 1989) have been developed during the present program of research (Saroyan, 1989; Duy, 1990; Rahilly, 1991). These existing schemes were used as a basis, and modified to answer this study's particular questions, while staying within the parameters of the program of research, and maintaining a firm theoretical base. Since the existing coding schemes are all based on the assumption that revision is a problem-solving task, where the reviser conducts a search through the problem space, Newell and Simon's (1972) problem space model was used as a basis for developing codes (see Chapter 2, especially Figure 1 for details). The codes will be elaborated in the next section.

One protocol from the "without data" condition and one from the "with data" condition were selected at random and coded using the previously developed coding scheme. Gaps in this scheme emerged, and new codes were developed to reflect the

requirements of the study, taking care not to be swayed by the idiosyncratic syntax of each subject. Seventy-two codes were necessary to account for all the types of segments in the protocols.

To assist in the coding process, and to make it easier to train coresearchers for checking coding reliability, a flow chart was developed (see Appendix C), and used as a job aid. Difficulties in coding caused by idiosyncratic speech patterns of the revisers were overcome by constant back reference to the problem space as the theoretical base for the coding scheme. In addition, a dendrogram (Krippendorff, 1980) was developed to cluster codes which were closely related and thus serve as a taxonomy (see Appendix D).

Coding of all tapes was done while listening to the audiotapes of the think-alouds, so that "non-linguistic data" (Jones, 1987) from the tapes could be used to aid the process. These data were often voice inflections, such as those denoting sarcastic comments which might otherwise have been wrongly coded.

Coding scheme

This coding scheme contained four major categories based on the stages of Newell and Simon's (1972) problem space model described in Chapter 2, namely, recognizing the existence of a problem, evocation or construction of a problem space, selection and application of operators, and evaluation of new knowledge states. The codes of interest in this research were those relating to selection and application of operators. To accommodate all the segments, it was necessary to add two other categories of codes: Theoretically irrelevant or conversational units, and codes representing actions taking place outside the revision problem space.

The coding scheme will now be described, using examples from the protocols. For each example, the following information is given: the code assigned to the segment; the subject's identifying code and the session number; the segment number; the segment. For example,

ES+ S4T1 145 I kind of like the style of that paragraph. //

represents a positive evaluation statement (ES+), made by subject S4 in the first session (T1). The segment number is 145, and the segment being coded is "I kind of like the style of that paragraph."

For some examples, nearby segments are quoted in order to clarify the example. References to Saroyan, Rahilly, Israeloff and Tremblay indicate previously developed codes. More information on these codes may be found in the relevant theses.

Recognizing the existence of a problem

Evaluation Statements (ES-, ES, ES+)

(Saroyan)

A segment which refers to the reviser's judgments about the text. The result of a comparison between the current state and the goal state. This category represents negative comments which do not explicitly state the source of the problem (ES-), neutral (ES) or positive (ES+) statements, and expressions of preference, judgement, internal feelings and observations expressed by the subject. When these statements are negative, they may subsequently lead to a problem identification or revision statement. When they are neutral or positive, they do not constitute a problem, since there is no difference between the observed and goal states.

Examples:

ES+ S4T1 ES+	145 146	I kind of like the style of that paragraph. // It's sort of getting me hooked a little bit. //
ES S5T2 ES		I'm not saying it's right or wrong, // I'm just saying it strikes me, that it's slightly different.//
ES- S7T1	460	I think Aach! This isn't worth fixing//

When an evaluation statement contains explicit reference to a problem with the instructional materials, it is coded more specifically as a problem identification, and is further coded to identify the data source initiating the problem.

Problem Identification (PI)

A segment which contains explicit reference to an observed problem with the instructional materials, and is initiated by the reviser.

Examples:

PI S4T1 131 Since my interest in this is sort of more as a consumer than anything else, I'd like to know why the stuff -- why cyclamate has been banned. //

Problem initiated by data (PIS, PIA, PIL, PIT, PID) (Le Maistre)

A segment which contains explicit reference to an observed problem with the instructional materials and is initiated by a comment from the subject-matter expert (PIS); from the target population expert (PIA); from the learners (PIL); by the test scores (PIT); or by more than one source (PID).

Examples:

PIA	S1T1	1370	Well, I think that's a little too strong; but, it's mostly I agree. //
	S6T1	799	Now, she's getting at what they need to know. What would they want to know. //
PIS		800	And she's in a roundabout way, saying that this isn't useful information. //
PIL	S4T2	718	They don't understand trans. //
ΡΙΤ	S4T2	628	But, in terms of making changes based on post-test, there are some things the fact that the questions that are a little more indirect or a little more conceptual, were badly there was lower rate of success on on those questions, leads me to believe that, on the one hand, students are reading this from from memory. //
חום	S7T2	1675	because I've just noticed that the experts didn't like that

PID S7T2 1675 because I've just noticed that the experts didn't like ... that last paragraph on page 4 //

Problem corroborated by data (PCS, PCA, PCL, PCT, PCD)

(Le Maistre)

A segment which contains explicit reference to an observed problem with the

instructional materials and is corroborated when the reviser subsequently reads a comment

from the subject-matter expert (PCS); from the target population expert (PCA); from the learners (PCL); by the test scores (PCT); or from more than one source (PCD). Examples: '

PCA	S7T2	281	Hey, that's what I thought too when I first read it! //
PCS	S5T1	746	And I think the subject-matter expert refers to it, // (this was the only example of this code)
PCL	S1T1	178	they questioned that word being there too//
PCT	S5T1	1047	I think that is what I was talking about, the whole problem of how it's presented as very U.S. and people aren't //

PCD (there were no examples of this code)

Problem identified despite data (PDS, PDA, PDL, PDT, PDD) (Le Maistre)

A segment which contains explicit reference to an observed problem with the instructional materials and is identified by the reviser in contradiction to a comment from the subject-matter expert (PDS); from the target population expert (PDA); from the learners (PDL); by the test scores (PDT); or by more than one source (PDD).

Examples:

1	PDA	S1T1	904	I think they're going to get bored if it's presented this way. / /
	PDL	S7T2	1008	well, they didn't pick out the fact that there was three in there. //
	PDS	S6T1	785 786	And she says, "it's a very important point"; // but, nobody understands what it means! //
	PDT	S3T2	57 4	it's quite a high percentage of people got the answer to the question even though it is very, very dense. 79% // (this was the only example of this code)
	PDD	S7T2	646.	Oh, well, nobody else had problems with that, just me. //

Problem rejecting data (PRS, PRA, PRL, PRT, PRD)

A problem identified in the data is rejected by the reviser. The rejected problem may be identified by the subject-matter expert (PRS); by the target population expert (PRA); by the learners (PRL); by a test score (PRT); or by more than one source (PRD). Examples:

PRA S3T2 1001 I don't know why he's isolated selenium and not beta carotene, unless we're really going into details on what all those things are...//

PRS S5T1 763 I think that's a gratuitous comment, in this case. //

PRL S1T1 497 I'm not sure that this sentence in unclear, //

PRT (there were no examples of this code)

PRD (there were no examples of this code)

Problem elaboration (PE)

(Rahilly, revised by Le Maistre)

A segment enlarging on an identified problem. These are always associated with a problem identification statement; usually following it, but occasionally preceding it, and acting as a lead-in to the actual problem statement.

Examples:

ΡI	S8T2	525	All of this stuff, I am questioning on it. //
PE		526	I don't know what it is doing there, I guess, basically. //
PE		527	I can't assimilate that in any meaningful way at all. //

Problem reiteration (PR)

(Rahilly)

A restatement of a previously identified problem, usually, but not always in the same words as the problem identification.

Examples:

PI	S5T1	321	I'm sitting here in 1992, going so, what happened? //
PR		330	Because, I'm saying, okay, this is already old; //
PR		332	In the first paragraph, that sets me off, that this is out of date. //

Task representation (TR)

A segment expressing the subject's setting up of a short term goal, or a progress report on the subject's current actions. Thus, this code is future-oriented or presentoriented.

Examples:

TR	S7T2	109	Okay, let's see what they were supposed to learn. (reads pretest/posttest questions)//
TR	S7T2	1109	Let's assume that I'm going to implement these changes afterwards, okay. //

Task clarification (TC)

(Le Maistre)

A comment or question addressed to the administrator to check understanding of the task, or to clarify the parameters of the task.

Examples:

TC	S4T1	26	What other support do they have available besides the materials?//
TC	S2T1	344	So the comments I have, there are no comments from real people. //

Text talk (TX)

(Saroyan, revised by Le Maistre)

A comment representing the subject's paraphrases or summaries of the text of the stimulus materials, or knowledge that the reviser has acquired from the text of the stimulus materials.

Examples:

TX

S5T1 798 So, I understand that for a child with PKU, that this is a serious concern. //

(coded as Task Talk by Saroyan)

Verbatim statements (VS, VSS, VSA, VSL, VST, VSD)(Saroyan revised by Le Maistre)

63

This code is included in this stage of the problem solving sequence because when a verbatim statement precedes or immediately follows a revision, it enables the revision to be linked to a source.

A verbatim statement is one which is read by the subject and quotes the text, usually, but not exclusively, verbatim, A paraphrase also identifies the source of the statement, because occasionally the subject may paraphrase a comment from the feedback data, or comment on the learners' success on a test item without reading the score on the item. The subject may be reading the module (VS); a comment by the subject-matter expert (VSS); a comment by the target population expert (VSA); a comment by the learners (VSL); test information (VST); or reading information about how feedback data were collected (VSD). Reading about how feedback data were collected provides subjects with information on the validity of data sources and allows them to assess the credibility of the source.

Examples:

•	vs	S6T1	310	So, "saccharin costs only one twentieth as much as sugar."//
	VSA PIA	S1T1		"Tables are always troublesome."// They sure are, especially where they're badly done. //
	VSS	S3T2	111	Okay, so subject-matter expert is familiar with that report//
	VSS		112	And he didn't agree with it. //
	VSL	S4T2	930	Students say, "awkward wording, explain meaning." //
	VST	S6T1	1133	The first one didn't get very high test scores on saccharin, for that first question.//
	VSD	S4T2	13	So the SME has a "BSc Nutrition, ta, ta, ta, ta. Years of experience, thirteen years, stated area of expertise. And your content, three; audience one." //

Knowledge statements (KS)

(Sarovan, revised by Le Maistre)

A segment reporting statements of knowledge, or absence of knowledge. The reviser may propose a reason for making the revision, based on his own procedural knowledge of revision, or declarative knowledge of audience or revision. This code is included in this section because when a reviser justifies a revision based on previous knowledge he is evoking an already used problem space from long term memory.

Exam	ples:
	P100.

KS	S5T1	217	Because overall statistics like that, very often, people just distance themselves from them. $\ensuremath{/\!/}$
KS	S6T1		but then how you'd put it together so that the important parts are memorable for these people is more to do with the strategy, I think, //

Selection and application of operators

This stage of the problem-solving model contains the act of revision. Codes for revision statements parallel the codes for problem identification. Sometimes revision codes are preceded by a problem identification, sometimes problem identification codes are implicit. If subjects have defined the task for themselves as "improve the materials", it may not be necessary for them to identify specific problems overtly before searching the problem space, especially since it is characteristic of experienced performers to omit overt references to problem definition during problem solving. Like problem identification codes, revisions are coded to identify the source prompting the revision, and this set of codes forms the basis for analyzing the results.

Revision statement (RI)

(Saroyan revised by Le Maistre)

A segment which proposes a revision to the instructional materials and is initiated by the reviser. This is the only possible revision statement when no feedback data are provided.

Examples:

- RI S8T1
- 165 So, basically, this information needs to be put in some kind of a table form; //

Revision initiated by data (RIS, RIA, RIL, RIT, RID) (Le Maistre)

A segment which proposes a revision to the instructional materials and is initiated

by a comment from the subject-matter expert (RIS); from the target population expert

(RIA); from the learners (RIL); by the test scores (RIT); or by more than one source (RID).

Examples:

•	RIS	S3T2	1268	Well, if it is controversial research and we believe our subject-matter expert, I suppose we should mention that some experts find it controversial. (laughs) //
	RIA	S8T2	883	Perhaps. I see her point, that the research could be told in a more of a story form. //
	RIL	S4T2	641	Where they say, define terms, I would make sure that those terms are defined, of course, //
	rit Rit	S1⊤1	756 768 769	And question ten, they don't have the right answer here.// And it seems to me that this could be said a little bit shorter, // and a little bit more punchy. //
	RID	S7T2	1800 1802	<u>don't use unneccessary terms</u> // That bothered everybody.//

Revision corroborated by data (RCS, RCA, RCL, RCT, RCD)

(Le Maistre)

A segment which corroborates a revision when the subject reads a comment subsequent to proposing a revision. The corroboration may come from the subject-matter expert (RCS); from the target population expert (RCA); from the learners (RCL); by the test scores (RCT); or by more than one source (RCD).

Examples:

RCA :	S3T2	408	Okay, so that sort of backs up my request for a clearer definition of the term "environmental."//
RCS	S8T2	276	That's a, That's a good idea. //

RCL S7T2 2006 Well, we've covered that with the sidebar. //

RCT S4T2 526 So, that just confirms that we have to fix that table. //

RCD (there were no examples of this code)

Revision proposed despite data (RDS, RDA, RDL, RDT, RDD) (Le Maistre)

A segment which proposes a revision to the instructional materials and is made by the reviser in contradiction to a comment from the subject-matter expert (RDS); from the target population expert (RDA); from the learners (RDL); or by the test scores (RDT); or from more than one source (RDD).

Examples:

	•		. ,
RDA	S6T1	1097	So somehow they've got to be built in in an interesting way,//
RDA		1098	or just don't use the words.//
RDA		1099	But I wouldn't make a what do you call a lexicon.//
VSL	S1T1	374	"this sentence should come before 'for example'." //
RDL		375	I think that the sentence before that could say, let's look at this a little, in a little more fine detail. $\prime\prime$

RDT (there were no examples of this code)

RDS (there were no examples of this code)

RDD (there were no examples of this code)

Revision data rejected (RRS, RRA, RRL, RRT, RRD)

(Le Maistre)

A revision proposed in the feedback data is rejected by the reviser. The rejected revision may be proposed by the subject-matter expert (RRS); by the target population expert (RRA); by the learners (RRL); by test data (RRT); or by more than one source (RRD).

Examples:

RRS S5T1 8

898 Yes. I don't like cartoons of fat ladies in lines with black forest cake and a diet soft drinks. Cartoons making fun of people. (Subject-matter expert had suggested a cartoon in a comment which the subject did not read aloud)// RRL S4T2 415 My God, I don't believe somebody said that. //

RRT (there were no examples of this code)

RRD (there were no examples of this code)

Revision elaboration (RE)

(Le Maistre)

A segment enlarging on a proposed revision. These are always associated with a revision statement; usually following it, but occasionally preceding it and acting as a lead-in to the actual revision.

Examples:

RI	S8T1	432	one of the ways that this could be played with, is, you
			know, more in the format of a, "did you know that." //
RE		433	Pulling out some of those surprising research results. //

Revision reiteration (RR)

(Le Maistre)

(Le Maistre)

A restatement of a previously stated revision, usually, but not always in exactly the same words as the original statement.

Examples:

RR S4T1 346. We already wrote in the margin, a little earlier, the issue of ethics of animal research. //

Solution planning (SP)

A segment in which the reviser proposes a revision tied to the current task which he may not intend to carry out. Sometimes contain "would", "could", or expressions of hypothetical reasoning. Solution planning is also a characteristic grammar of subjects who

demonstrate tentative speech patterns and tend to make suggestions rather than assertions.

Examples: SP	S5T2	634	And I would ask a competent graphics person to make some suggestions as to how to do that. //
SP	S6T1	535	so I think that they should start with asking the the participants or the people who who thought this course would be a good idea, in the first place, what what was it that they wanted them to focus on, out of all this. // we could leave out a lot of stuff.//

Strategy statements (ST)

(Saroyan)

A segment referring to the reviser's usual procedure, but not tied to the current task. This code usually contains words such as "would", "generally", "usually." Strategy statements are included in selection of operators, because they are based on the reviser's strategic knowledge. The reviser is calling up a task which he perceives to be isomorphic, in an attempt to reuse a previous problem space (Newell & Simon, 1972, p. 851)

Examples:

ST ST ST ST ST	S2T1	136 139 140 141 142 143	You know what I would do. // What I would do right now, if I were at the office,// I would stop reading, // I would get on the phone // and I would call the subject-matter expert, // and I would ask him, "what the hell is it that you want to do"? //
ST	S1T1	962	Normally, I would like to have that before I start //
ST	S1T1	1043	Again, I'm I'm very big on telling people where they are and where they're going.//

Evaluation of new knowledge state

Solution evaluation (SE)

(Le Maistre)

The subject assesses the effectiveness of the revision he has made, and either

accepts it or rejects it.

Examples:

SES1T2263Again, my headings.Well, okay, it still fits. //SE264That would -- I guess it would be okay. //

Theoretically irrelevant or conversational units

<u>Task talk (TT)</u>

Comments or questions addressed to the administrator concerning the research task and its methodology.

Examples:

TT S3T1 133 Okay but did you want me to write down all the changes that I'm suggesting?//

Dialogue (D)

<u>(Saroyan)</u>

A segment relating to any conversation with the administrator that was not intended to clarify the task, nor discuss the research task, nor describe a location.

Examples:

D	S7T1	435	you can't edit on the computer, //
D		436	because the problem with the computer is, it doesn't give you the whole. //
D		437	It only gives you a little discreet chunks; //
D		438	and it's very hard to see the whole. //
D		439	So, you need the whole, you need the hard copy. //
D		440	But, when you get down to moving text around, whoa! //

Administrator's responses (K, T)

(Le Maistre)

Any comments made by the researcher, including prompting words, answers to questions of clarification, reinforcing comments, dialogue, checks on location of comments.

Examples:

к	S7T2	102	So, in this case, first year undergraduates. That sort level, anyway.//	of
т	S6T1	96	But, those are the results of this test. Not the test that came with the course.//	ıt

(Israeloff/Tremblay)

Locating statements (L)

A segment in which the subject identifies his location in the text, often in the form of a sentence fragment or phrase, sometimes in response to a prompt from the administrator.

Examples:

L S6T1 218 So now, on the next page,//

False start (FS)

A segment which does not represent a complete thought.

Examples:

FS	S3T1	442	And //
FS		444	I suppose. //
FS		445	Oh, no, actually, it could be //
FS		446	it could be //

Boundary markers (BM)

Verbal or non-verbal utterances forming a break or link between segments.

Examples: "Um; OK; Hmmm; Let's see; Well"

Unrelated talk (UT)

JUL

Any statement made by the subject which does not relate directly to the revision

task.

Examples:

S2T1 208 Guess what. I went on a cruise this summer and didn't take any sun.//

Codes representing actions taking place outside the revision problem space Monitoring (M) (Le Maistre)

Self-referential meta-comments reflecting the subject's report or analysis of his previous thoughts or actions. This code appears to be similar to the task representation code, but is <u>past-oriented</u>, rather than <u>future-oriented</u>. It reflects executive control of the

(Sarovan)

(Sarovan)

(Sarovan)

process, and is outside the revision problem space. These comments represent the subject stepping back from the task and evaluating the solution <u>process</u> -- as opposed to evaluating the <u>solution</u> (SE).

Examples:

М	S1T1	235	I'm, again I'm jumping all over the place in terms ofYou know, grosso, you know big plan, little plan stuff but
М	S4T2	584	Right, now, I'm trying to avoid getting into the content, is what I'm doing.//
M M	S7T2		(laughs) I just realized that I'm running into detail // and I should maybe stop and think of (pause) structure. /

These examples explain the codes necessary to code completely the think-aloud protocols for each subject. In summary, 72 codes, developed according to Newell and Simon's (1972) theory of human problem solving, were necessary for coding the verbal protocols.

Reliability and stability. After the whole data set had been coded by the researcher over a period of about 6 months, it was recoded by the researcher to check for stability. This was repeated for randomly selected transcripts over a further 4 month period until the stability was at least 90%.

Four coresearchers were given a brief training period in the coding scheme and provided with the flowchart shown in Appendix C which represents the coding process. Each was asked to code a total of 273 segments taken from the middle of 4 different revision sessions, when it was assumed that the subjects' performance had stabilized and before the subjects had become tired. Overall, at least two coders agreed with the researcher's coding 78% of the time. When only revision linked codes were counted, at least two of the coders agreed with the researcher's coding 84% of the time. Discrepancies in the coding were negotiated between the researcher and the coresearchers. In the few cases where discrepancies could not be resolved, the researcher's decision was taken as the standard. Reduction of the number of codes obtained from the data sets. Each of the 16 data sets was segmented and coded into the 72 categories described above, and tabulated, as shown in Appendix E. These categories may be grouped into two main classes: those which relate to the human problem-solving model, which are therefore theoretically significant, shown in Appendix F; those which are not relevant to the human problem-solving model and which are theoretically insignificant. Theoretically significant codes may be further grouped into two categories: those related to revisions, shown in Appendix G, and those not related to revisions. This classification is summarized in Figure 7.

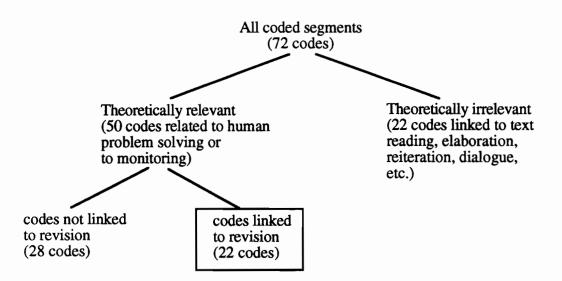


Figure 7. Classification of coded segments to isolate revision segments

Since the focus of the study was the subjects' revision performance, codes related to revision needed to be isolated from the data set. This was done in two stages. The first stage was the removal of theoretically irrelevant segments, which left codes related to the task at hand, but still included segments not related to revision. The second stage was the removal of segments not related to revision. The theoretically insignificant statements were:

 All locating statements (L) and verbatim statements (VS, VSS, VSA, VSL, VST, VSD) used to tie comments made by the subjects to particular sources, and no longer necessary after that link had been made;

2. Text talk (TX), important to understanding the reviser's representation of the materials, but irrelevant to counting the revision statements;

3. Problem elaborations and reiterations (PE, PR), and revision elaborations and reiterations (RE, RR), whose inclusion would have inflated the numbers of problems identified and revisions made;

4. Administrator's responses (K, T);

5. All theoretically insignificant dialogue, comprising general evaluation comments on the materials (ES+, ES), dialogue (D), unrelated talk (UT), false starts (FS), and boundary markers (BM).

Removing these segments left 50 theoretically relevant codes, shown in Appendix F, which were grouped under the four headings of the human problem-solving model as follows:

<u>1. Recognizing the existence of a problem:</u> negative evaluations (ES-); all problem codes (all remaining codes beginning with P-).

2. Evocation or construction of a problem space: task representation and clarification (TR, TC); knowledge statements (KS).

3. Selection and application of operators: solution planning (SP); strategy statements (ST); all revision statements (all remaining codes beginning with R-).

4. Evaluation of new knowledge state: solution evaluation (SE)

Finally, monitoring (M), although it was an action taking place outside the problem space, provided valuable information on the subjects' assessment of the task they were performing. It was deemed to provide information as important as that provided by knowledge statements and strategy talk, and was included in the theoretically important codes.

The 50 segments bearing theoretically relevant codes were tabulated, but still contained material irrelevant to the present study. The stage of the human problem-solving model of most interest to this study is the selection and application of operators, so segments relating to the other three stages were not included in the analysis. Codes relevant to the selection and application of operators were: all revision codes (all codes beginning with R-), solution planning (SP) and strategy statements (ST). Strategy statements report the subject's actions in earlier problems, so they do not produce revisions tied to the module currently being revised by the subject. For this reason, strategy statements were not counted in further analysis. This left for analysis all segments given revision related codes (R-) and those related to solution planning (SP), making 22 codes in all.

The revision codes isolated in this way were closely related to the sources of data giving rise to the revision. The first three questions of the study address different components of these sources, so the 22 codes linked to revision were grouped according to the requirements of the particular question being addressed. This will be described in detail in the next chapter. Mean values were calculated where relevant, and analyses of variance were performed.

<u>Other sources of data for analysis</u>. As well as the transcribed, segmented, and coded protocols from the think-aloud, other data were collected to investigate the actions of the eight revisers.

Written comments

It has been observed that problem solvers place reliance on external representations during problem solving (Newell & Simon, 1972; Olson & Biolsi, 1991). Since external representations provide an auxiliary display for the information being attended in short term memory, they may be imagined to move the subject through the problem space more efficiently. The written trace provided by the reviser may therefore provide additional information about his passage through the problem space, unless the subject says aloud what is being written. In this case, the written record would merely duplicate the verbal data.

Some subjects wrote extensively, whether on the materials or on scratch pads, while they were revising. These written comments could be classified as: the actual revisions being proposed or effected; notes to inform later actions; notes summarizing what had been learned from the modules; sections of the modules which were highlighted with colored pens.

In all cases, subjects read aloud what they were writing, so it was not necessary to include the written comments as an extra data set. However, examination of the written comments often helped to clarify or locate references in the oral comments while these were being transcribed, segmented and coded.

Retrospective debriefing

The retrospective debriefing was also transcribed for each subject. These transcripts provided clarification of issues not clear to the administrator, elaboration of points made during the think-aloud, and information about the revisers' training, experience, usual practice, reasons for actions during the task, and perceptions of the task. The retrospective debriefing provided data for answering the fourth and fifth questions, and will be quoted extensively in the next chapter.

Summary

This chapter has described the design and method followed in this study. The order of administration of both the module and the feedback condition were counterbalanced. This allowed investigation of the main effect, the provision or lack of feedback, as well as a check on whether the modules were of equivalent quality and on whether there was any order effect of treatment. Obtaining feedback data on the two modules added an extra layer to the collection of data. The data collected for three masters' theses were used to provide review data from experts and learners on the first module, and the methods of these theses were replicated to obtain feedback data on the second module. These data sets were summarized as part of the materials package given to each reviser.

Each of 8 subjects (S1 to S8) revised the printed instructional material during two sessions (T1 and T2), while thinking aloud. This produced 16 data sets (S1T1 through S8T2). In each of the two sessions, subjects were given either Module 1, "The diet-cancer relationship", or Module 2, "Artificial sweeteners" and were asked to revise it, either with feedback or without feedback. Subjects' oral comments were tape-recorded and their written comments were collected. After each session, there was a retrospective debriefing to clarify and elaborate on the oral comments, and to obtain demographic information.

Verbal protocols from the think-aloud sessions were transcribed, segmented and coded. A coding scheme was developed which allowed extraction of the revision statements made by subjects and the identification of the source of the revision, whether from one of the feedback data sources, or from the reviser's own knowledge. Comments made during the retrospective debriefing were also transcribed. The protocols produced by the think-aloud procedure were the primary data set, but the methods used for the study provided a variety of data for analysis. The results of these analyses will now be described.

CHAPTER 4: RESULTS

This study investigated five questions relevant to the procedures followed by instructional design practitioners when they were revising. Eight instructional designers thought aloud while revising two modules, one module with the type of feedback data recommended in the standard model of instructional design, and one module with no data. The subjects' protocols were audiotaped, transcribed, segmented and coded according to a coding scheme based on Newell and Simon's (1972) model of problem solving. The segments coded as revisions were extracted from the data set, and analyzed using standard descriptive and inferential statistical procedures, which will be described in this chapter.

This chapter opens with a demographic description of the subjects, obtained during the retrospective debriefing, and continues with an overview of their performance of the research task. This overview will include the subjects' definition of the task they were performing, a description of their facility with the think-aloud technique, and a narrative description of their actions while they were revising. It will continue with a reprise of the results of isolating the coded segments of interest to the study and then present the results pertaining to each research question, using the questions as a framework.

Comparing the revision segments produced in the "with feedback" and "without feedback" conditions allowed the first three research questions to be addressed, namely:

1. How much personal knowledge will revisers incorporate, if they are provided with verbal data from subject-matter experts and audience experts, verbal comments from learners, and test data from learners?

2. What attention will instructional designers give to each data source, if they are provided with verbal data from subject-matter experts and audience experts, verbal comments from learners, and test data from learners, and asked to revise print materials?

3. Do revisers provided with external feedback incorporate as many of their own revision suggestions as revisers provided with no external feedback?

Analysis of think-aloud segments not related to revision, but classed as theoretically relevant, and also of the comments made during the retrospective interviews produced some answers to the fourth question:

4. How do experienced revisers establish priorities among data sources?
Finally, the performance of the subjects was examined to provide information for the fifth question:

5. When they revise materials, do instructional designers follow the standard model presented during their training?

The chapter will end with a summary of the results, and the results will be discussed in Chapter 5

Description of Subjects

<u>Qualifications</u>

All subjects had completed postgraduate courses in educational technology. One subject had a PhD, five had completed Master's degrees, and two had completed coursework for a Master's degree, but had not finished the thesis requirement. Seven of the eight subjects had attended the same University. All subjects had taken at least one course in Instructional Design, three had taken two courses and one had taken three courses. Five subjects reported that they had taken no courses specific to formative evaluation or revision, and seemed surprised by the question's implication that such courses existed, but 3 reported having taken one course in formative evaluation. The subjects' experience in instructional design ranged from 5 to 18 years, with a mean of 9.6 years and a median of 8.5 years. Chi-square tests performed on these data as part of another study showed that the subjects were demographically similar (Bordonaro, 1993).

Preferred Reference Material

Subjects were asked which instructional design textbooks they had used during their studies. All responded with at least one textbook, which they referenced by the name of the author, rather than the title or other publication information. The texts and the number of times each was mentioned are represented in this way, without other citation, in Table 4.

Table 4.

Textbooks Used by Subjects During Training

Textbook	Number of respondents
Dick & Carey	5
Gagné	1
Gagné & Briggs	1
Mager	1
Romiszowski	1
Rossett	1
Stolovitch	1

Since half of the subjects had trained eight or more years ago, they were then asked which textbooks they still consulted. Their responses are shown in Table 5.

Table 5.

Textbooks Still Consulted by Subjects

Textbook	Number of respondents
Dick & Carey	5
Carlisle	1
Gagné, Briggs, & W	ager 1
Kemp	1
Powers	1
Reigeluth	1
Romiszowski	1
Rossett	1
Rothwell & Kazanas	1
Stolovitch	1
Wittrock	1

Some subjects reported a preference for journals over textbooks as references. All subjects reported that they subscribed to at least one journal relevant to instructional design, with two reporting four journals each, and a median value of three for the group.

Affiliations and Conferences

All subjects belonged to the local branch of the National Society for Performance and Instruction (NSPI), some also belonged to NSPI International. Four of the eight also belonged to at least one other relevant professional organization. Seven of the eight had attended at least one professional conference within the preceding two years.

Familiarity with Materials and with Audience

After each session, during the retrospective interview, subjects were asked to rate their familiarity with the content of the module they had been revising and with the intended audience for the materials, using a scale of 0 through 4. Familiarity with the content was rated from 0 to 2.5, with a mean of 1.7 across both modules, and a standard deviation of 0.81. There was less variation when subjects assessed their familiarity with the intended audience, which was rated from 2 to 3, with a mean of 2.6 across the group, and a standard deviation of 0.40.

Performance of Research Task

Task Definition

To allow subjects to revise under conditions which were as naturalistic as possible, the task was not defined closely for them; they were simply instructed to "make any necessary revisions." The ill-defined nature of revision would, in any case, have made it difficult to prescribe a definition, since in an ill-defined problem such as writing, task definition becomes elaborated during task performance (Breuleux, 1991). In terms of the problem-solving model, the subject's definition of his task would be the first step in setting up a problem space for the problem at hand, because "people only solve the problems they give themselves to solve" (Flower & Hayes, 1980, p.22). Self-definition of the problem, while allowing the subject to establish an appropriate problem space, introduces a potential difficulty for a research study if the subject's definition does not match the definition proposed for the study.

This study investigated the actions of the subjects during the act of revision, and care has been taken to distinguish revision from review or editing. In Chapter 2, review

was defined as the assessment of draft instructional materials, with the intention of subsequently improving them. Revision was defined as any changes, short of rewriting, that the reviser deems necessary to improve the materials. This includes editing, which was defined as the superficial level changes made during revision. It was important to check the subjects' perception of the task to ensure that their definitions matched those proposed for the study closely enough so that they could be said to be performing the task envisioned for them.

As outlined in Chapter 3, and shown in Appendix B, the subjects' definitions were checked twice during each session. First, after subjects had read the task description, they were asked to state the task as they understood it. Second, their perception of the task was checked by asking during the retrospective interview: "If you go back to the office and have to describe this task to a colleague, how would you describe the task you've just completed?" Responses were stable across the duration of the task. In the responses, the word "revise" occurred 16 times and "review" 4 times. Other terms used by the subjects in defining their task were: "make it more effective", which occurred twice, "make recommendations", "assisted re-write", "commenting", "suggest changes", "recommend changes", all of which seem synonymous with the study's definition of revision. No-one used nor implied the word "edit" during the description of the task. However, "edit" emerged in the think-alouds of three subjects, and therefore became an item for discussion during their retrospective interviews. This provided interesting insight into the subjects' definition of their task.

Subject S4, appeared to be using the word "edit" interchangably with "revise", but when asked to distinguish between them, was clear about the meaning of "edit." (In each case, "K" refers to the researcher).

K You said that you wouldn't do the grammatical changes, that you'd give -you'd ask the author to give those to someone who really had good command of English and writing and so on and so forth. And then you said -- later on you used the term edit a lot, "I'd edit this down", "I'd edit this out."

How are you using the term edit, there? And would you use the term edit in the grammatical changes?

- S4T1 Uh, no. When I was saying I would edit this down, I really meant just finding the overly technical content, or finding excessive instances -- you know, instances of excessive technical content and editing that out by just saying -- just taking it out.
- K Okay, so by "edit" you just mean "remove"?

S4T1 Yes.

Subject S1 was less able to explain the distinction he was making between editing

and revising.

- K How do you distinguish between editing and revising? (....)
- S1T2 (....) there's -- there's always the two components to it. One is, is I have to come up with some kind of an -- or, I feel obliged to put some kind of an instructional model with seven steps to it, or, seven elements, or something like that. To make it easier for people to understand. On top of that, I -- then, what I'll do, is I'll go through and I'll try to take sentence by sentence, or -- or, idea by idea, and plug it into that model. But, usually, I can't use those ideas as they were. First, of all, because they're badly written. And second of all, they won't fit in the -- in the model. It would be disjointed to try to put it back together there. I'm not sure I'm answering your question though.
- K So, which of those levels do you call editing? The taking of sentence by sentence part? Or, re-writing? (....)
- S1T2 It's a really good question. I guess, I go through first, and see if I can sense a model or a structure.
- K Which is what you actually did with both these pieces.
- S1T2 Yea. And then, I see if the material can be shuffled to fit that structure. And if it can't; and I feel that there's some validity to the way it's written and the way the flow of ideas there, then I'd go back and question my model and say, the model was no good.

Subject S7 spent several minutes distinguishing among editing, review and

revision:

- K Can you make a distinction for me between revising and editing? because you're using both terms. What's your definition of one or the other? or both?
- S7T1 My definition is probably idiosyncratic to me. When I'm asked to review something, I'm not editing it. I'm reviewing it for its instructional design and its fitness. I'm reviewing what I've got in front of me to offer my comments.

- K Where does revising fit into that? Are you using reviewing and revising...
- S7T1 No revising now would be going in and (pause) editing. Yea, that would be more on the editing side of things. If I was revising this, if somebody said: this is yours to work with, feel free ... I think think this is where my confusion's come from ... This is yours to work with, feel free. I don't have to worry about going back to the SME now, I don't have to think about those things now, I just have to go through this thing and say (Snort), and really re-write this thing from top to bottom. Because for me ... this turns around in too many circles. Editing ... I mean if you'd given me five pages. I could have edited five pages, but the other thing is I'm hung up because you said this is the shorter night, and I'm thinking: Wait a minute, this is a lot of editing, this is heavy editing to get through all this. Revising is more ... If I was revising it under those conditions that it's mine with to do what I want, I would do a heavy edit as well as a restructuring. "I would like you to review it" means I'm going to offer my comments. "I'm going to revise it' means that I'm going to restructure probably and do a little bit of editing. But when I looked at this, I would go back to the person and say: Oh this is major heavy duty revision here, major heavy duty editing. Yea, I guess I'm playing with ... It's funny you should ask me that. I guess I've never thought about the differences in the words I've been using. This is major revision to this, almost ...
- K So are you saying that revision is a larger project than editing?(...)
- S7T1 You can buy, hire, technical editors.
- K But you can't hire revisers?
- S7T1 You can hire revisers, but usually editing comes with revision.
- K Which one is a sub-set of the other?
- S7T1 Oh Jeez, I don't think one is a sub-set of the other. I think one ... You can play the editing ... Hm, interesting, interesting. Can we dismiss reviewing? Because we've got that one covered. We understand what that means.
- K Reviewing means looking at it and going back to the SME and saying...
- S7T1 I've reviewed it and these are my comments and these are what I recommend that you do. Revising is mine to do with what I want, and it's my job, that's what I've been asked to do, I'm going to revise this thing. OK, I'm going to come up with a phrase that I would use to see if this helps me. "I've got to revise this and there's a lot of editing in it." By editing, I talking more about word changes. Revision can include word changes. Editing.... I suppose if I was a journalist, editing would be revising, but because I come from an instructional design background, I tend to think of editing as a sub-set of revising. Did you get that on tape?
- (....) OK let's say, you've got a fictional novel that you've written, the editor would go through to make sure that all the grammar's correct, the periods and the commas are all in the same, the paragraphs follow, the links are made between paragraphs ... Well yes, that sort of stuff too. A reviser tends to look more, to go broader and look more at the structure of things. With an editor, I'm not going to change the structure. Though I suppose... Yea, it's this... I'm an editor, I'm a reviser, and there's a big grey zone in

the middle (drew a Venn diagram with an intersection of two sets) where there's an overlap. There's the same overlap between design and development.

During the think-aloud, these three subjects had seemed to be using three terms important to this study in a casual sense, but it was clear after probing them for a more precise definition that their understanding of the terms matched those defined for the study. It was also clear from the protocols that they (and the other 5 subjects) were implicitly representing "editing" as surface changes, "reviewing" as providing input for revision, and "revision" as all changes short of completely rewriting, as the words were defined for this study. The task definition of all the subjects was close enough to that proposed for the study so that subjects could all be said to be satisfying the constraints of the experiment.

Subjects' Ability with the Think-aloud

The think-aloud process was new for all subjects. Although several subjects commented that it was a tiring process, none reported that it had affected their revision procedure. Transcripts of the think-aloud protocols show that few subjects needed prompting, the number of prompts ranging from 1 to 18 with a median of 3.5, for sessions ranging from 1.5 hours to 3 hours long. Those subjects who needed most prompting tended to do so in the "with feedback" condition. Input from the administrators was more likely to be in answer to questions from the subject, or non-verbal expressions of encouragement, rather than verbal prompting. Only one subject seemed to be formulating complete sentences before speaking. This was verified by the fact that few of her segments were coded as false starts (FS).

Mean Time on Task for each Session

Subjects performed slightly faster during the second session than the first, or when they did not have any feedback data, rather than when they had data, or when they were revising Module 2, rather than Module 1. Mean times are given for each pair of conditions in Table 6.

Table 6.

Time On Task For Each Pair of Conditions

	Mean time	Standard deviation
Time 1	2.34 h	0.40
Time 2	2.25 h	0.89
With feedback	2.63 h	0.64
Without feedback	1.97 h	0.66
Diet-Cancer Relationship	2.50 h	0.65
Artificial Sweeteners	2.10 h	0.03

The differences observed here did not seem to be great enough to merit further analysis.

Overview of Subjects' Revision Procedures

In the performance of an ill-defined task, it would be surprising to find subjects following a single heuristic, and it was observed that the eight subjects did not all proceed in the same way. For example, differences occurred in the order in which they read feedback comments in the "with feedback" condition. Nevertheless, notes taken by the administrators during the sessions, and substantiated by examination of the transcripts, show many commonalities among the subjects. In general, all eight subjects could be characterized by a high degree of organization. They spread out papers to be able to see them all, then put aside any that were not relevant. They glanced through the whole text and read section titles, took notes or made concept maps and summaries to gain an overview of the content. They all asked questions after a few minutes of reading the text to review and clarify the task they were to perform. In spite of this careful organization, all subjects moved back and forth from surface level changes, such as typographical errors, to major structural changes, such as reorganizing paragraph structure, throughout the revision in an apparently random manner. They were willing to "satisfice" (Newell & Simon, 1972, p. 681) temporarily on a solution, and defer judgement on a solution evaluation (coded as SE) until they had finished the whole task, or a large section, then revisit changes and reassess what they had done.

All but one subject (S₂) read the feedback data and commented on the sources of data when deciding whether or not to use the information. It will be seen later in this chapter that S₂'s results are anomalous in several aspects. With regard to the revisions they made, all subjects looked for objectives and were uncomfortable that there were none.

All subjects were critical of the format and presentation of the modules. The first comment made by Sg was: "It doesn't **look** like instructional material." Most were critical of the way the print medium was being used, and said that the text was dense, heavy, intimidating, boring, needed livening, had the wrong tone, was dry, needed to be shortened, or, in the words of one, "edited down." This thought was often expressed early in the revision process:

S8T2	74	Just general impressions then of that first paragraph, is
		that it's too dense, both visually and content wise. //

or later, after the subject's initial interest had worn off:

S4T2 765 it's too heavy, the tone is wrong, that it's too technical, it's too -- too dry. It's getting boring. //

Even so, one subject firmly rejected the "find the hidden killer" idea, which was suggested by one audience expert, as too sensational for the tone of the material. One added the comment that university students were probably used to this type of document.

Subjects suggested restructuring the modules, or actually tried to do this. The word structure was used often, whether by subjects trying to grasp the structure of the existing materials or trying to impose a new structure to make the text more instructional. Several subjects tried to make links between sections, or to improve the existing links. One proposed strategy suggested using successive levels of detail, starting off with broad sweeps and becoming more and more detailed at each level. Several subjects proposed starting the module with an advance organizer or pre-organizer, and concluding with a post-organizer. Thus the themes of organization and structure were important to all subjects.

Subjects suggested teaching strategies such as games, debates, experiments, a discovery approach, or a more interactive approach which personalized the materials for learners. Two subjects suggested text mapping, as a way of organizing the information for learners, one suggested that learners be taught concept mapping, as a way of facilitating their learning. Other suggestions were more student involvement, built-in exercises, embedded questions, and crossword puzzle to test for learning of terms. The technical vocabulary caused several subjects to suggest the use of a glossary, more definitions, or a pronunciation guide.

Among the major changes were suggestions to change the format completely to involve a portable job-aid to use while shopping, then supplementary reading for those learners who were interested, or to scrap the text completely and use a comic book format. Subjects suggested a two-column format, shorter paragraphs, with one idea per paragraph, illustrations, bulleted lists, headings and sub-headings, and sections organized in parallel form. Among the other formatting ideas were sidebars or boxes as drop-ins to the text, where interesting information not crucial to understanding the text could be inserted for

interest. This would separate the "need to know" from the "nice to know" and would help the learning process.

All subjects cut out phrases and words like "It should be noted" and "accordingly", and all noted surface errors such as typographical errors and faulty use of commas and hyphenation. All criticized the tables, and the revisers who had defined the overall objective at the level of familiarity proposed taking out the chemical formulas at the end of the "Artificial sweeteners" module.

In general, revisions tended to be those categorized as Instructional Design concerns by McAlpine and Weston (1994). The eight subjects demonstrated both similarity in their approach to the problem of revising, and confidence in their performance of the task.

Results of Segmenting the Transcribed Think-aloud Data

The think-aloud protocols were transcribed, segmented and coded, as described in Chapter 3, producing 18511 segments coded into 72 categories. The complete set of results is shown in Appendix E. The dependent variables under investigation were the number of revision segments and the different sources of feedback with which the revisions could be associated. The revision segments were isolated by removing all other segments from consideration, and the frequencies of the revision segments were tabulated. When theoretically irrelevant codes, such as conversational or para-verbal comments, were removed from this list, 6338 segments remained, coded into 50 categories, as shown in Appendix F. Finally, when only revision related codes were considered, by removing all segments not directly related to revision, there were 1981 segments, coded into 22 categories, as shown in Appendix G. This is summarized in Table 7, and represented graphically in Figure 8.

The fourth column of Table 7, obtained by dividing the number of theoretically significant segments by the total number of segments, shows that the percentage of relevant

segments ranges from 28.4% to 44.2%, with a mean of 34.8% and a standard deviation of 4.23 for all subjects. In other words, an average of about 65% of the comments made during the think-alouds were irrelevant to the task. Of the theoretically significant segments, the last column shows that the percentage of revision segments is more spread out, ranging from 20.6% to 52.5%, with a mean of 32.5% and a standard deviation of 9.48.

Table 7.

Subject Total segments		Theoretically relevant segments	Theoretically relevant as percentage of total segmen	Revisions	Revisions as percentage of theoretically relevant	
	(all 72 codes)	(50 codes)	(%)	(22 codes)	(%)	
S1T1	1527	582	38.1	180	31.0	
S1T2	510	199	39.0	72	36.2	
S2T1	1322	485	36.7	119	24.5	
S2T2	1028	347	33.8	163	47.0	
S3T1	618	246	39.9	83	33.7	
S3T2	1736	582	33.6	249	42.8	
S4T1	1002	443	44.2	102	23.0	
S4T2	1111	428	38.5	112	26.2	
S5T1	1540	47 1	30.6	127	27.0	
S5T2	696	235	33.8	77	32.8	
S6T1	1617	496	30.7	127	25.6	
S6T2	859	244	28.4	68	27.9	
S7T1	1080	350	32.4	88	25.1	
S7T2	2120	641	30.2	132	20.6	
S8T1	812	278	34.2	146	52.5	
S8T2	933	311	33.3	136	43.7	

Summary of Coded Segments for Each Subject

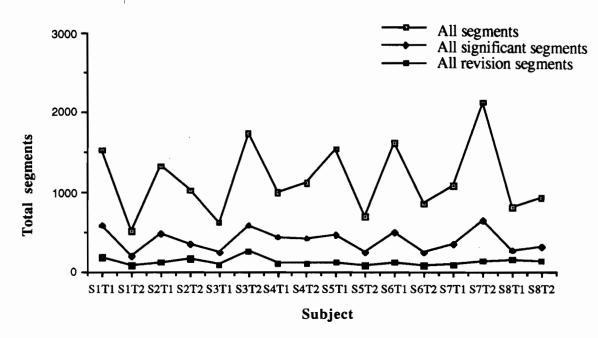


Figure 8. Comparison of the number of revision statements, and theoretically relevant segments with all coded segments

Grouping of Revision Codes

Examination of the transcripts provides evidence that, for all subjects, and in both levels of the feedback condition, the decision to plan a solution, rather than to carry it out, was predicated by two factors. One was the constraint of the research situation (whether lack of time, lack of access to alternative resource materials, or lack of access to a word processor). The other factor commonly causing a segment to be coded as a planned revision (SP), rather than a revision which was effected (RI), was the diffident speech mannerism of some subjects. Such subjects would say: "You could change...", rather than "I'm going to change...", but would continue the task as if the change had, in fact, been made. Solution planning was therefore included with revisions initiated by the reviser and counted as one code (RI and SP). In the condition without feedback (T1[S3, S4, S7, S8] and T2[S1, S2, S5, S6]), when there was no possibility of any revisions based on

external feedback data, only two codes were available to classify all the revisions: RI, revisions initiated by the reviser, and SP, solution planning.

After combining RI + SP, the revision codes in the condition with feedback $(T_1[S_1, S_2, S_5, S_6] \text{ and } T_2[S_3, S_4, S_7, S_8]$)were grouped in three different ways, depending on the research question being addressed. This will be elaborated for each of the three relevant questions, questions one through three.

Question 1: How Much Personal Knowledge Do Revisers Incorporate, if They Are Provided With Feedback Data From Experts and Learners?

When subjects were provided with feedback data, revision statements in the protocols were coded to link them to each of the four data sources (subject-matter expert, target population expert, learner comments, learner test scores) and also to the reviser's own personal knowledge. This was done by identifying a verbatim reading of a comment from a particular source before the revision statement, or by the reviser's reference to the source being attended. In some cases, the situation was clearer still: subjects performed the revisions even before reading the feedback data, so that the only source of information was the reviser's own knowledge.

Since the task was open ended, revisers were not limited as to how many revisions were either appropriate or necessary, so the actual number of revisions, rather than percentages, was counted.

The first categorization was designed to answer question 1, and is described in Figure 9.

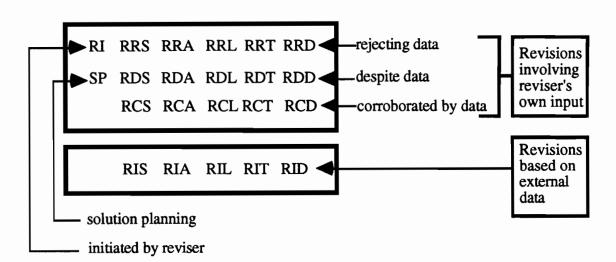


Figure 9. Revision codes linked to reviser's input or to external feedback.

The larger box in Figure 9 contains revision codes involving the reviser's own input (RI) and solution planning (SP), grouped with revision codes rejecting data (RR-), those despite data (RD-), and those corroborated by data (RC-). Segments where revisions were carried out (or proposed) rejecting data (RR-), or despite data (RD-) were classified as initiated by the reviser, because the reviser was using his own knowledge to override the information given to him. The smaller box groups all other revision codes which could be linked to feedback data (RI-) from one or more of the data sources given to the subject, and which were initiated by that data source.

The values for these two groups of codes are shown in Table 8, and represented graphically in Figure 10.

Table 9 summarizes the mean value for the number of revision statements linked to the reviser's own input and the mean value for the number of revisions from all other sources for all eight revisers.

Table 8

Subject /Time	Reviser	's own input	All other sources		
	Number	Percentage	Number	Percentage	
S ₁ T ₁	142	78.9	38	21.1	
S ₂ T ₁	118	99.2	1	0.8	
S3T2	222	89.2	27	10.8	
S4T2	70	62.5	42	37.5	
S5T1	86	67.7	41	32.3	
S6T1	115	90.5	12	9.5	
S7T2	98	74.2	34	25.8	
S8T2	79	58.1	57	41.9	

Number and Percentage of Revision Statements Linked to Reviser's Input and to All Other Feedback Data Sources for Each Subject

Table 9

Comparison of the Number of Revision Statements Linked to the Reviser's Input with Input from All Other Sources

Source of revision	Mean	Standard deviation
Reviser's personal knowledge	116.25	48.7
Total of all other sources	31.50	17.8

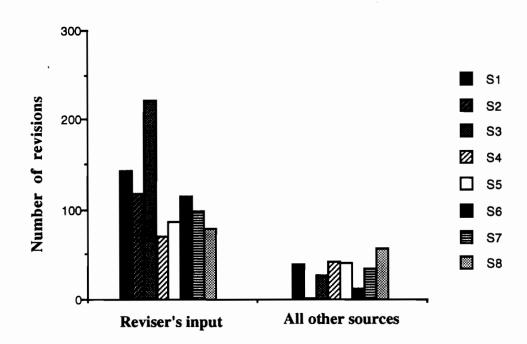


Figure 10. Number of revision statements linked to reviser's input and to all other feedback data sources for each subject.

A one-way repeated measures analysis of variance performed on the mean number of revision statements linked to each revisers' input, compared with the mean of the total number of revision statements linked to all other sources for each reviser showed that there were no significant differences for time of presentation or module. This allowed the results for individual revisers to be aggregated across these variables. The results of this analysis are shown in Appendix H.

The ANOVA for the within subjects analysis showed that there was a significant difference between revision statements linked to the revisers' personal knowledge and revision statements linked to the sum of all other sources, with significantly more revision statements (almost 79%) based on the reviser's own input. There was no effect caused by the time of administering feedback (session 1 or session 2), nor by any difference between the modules. There was no effect caused by the time of administering the feedback, nor by

the use of two different modules, nor any three-way interaction. This is shown in Appendix I.

Thus, even when revisers are provided with feedback data from the sources recommended in the standard model of instructional design, they use significantly more of their own input than input from the feedback data.

Question 2: What Attention Will Instructional Designers Give to Each Data Source If They Are Provided With Verbal Comments From Subject-matter experts and Audience Experts, Verbal Data and Test Data From Learners, And Asked to Revise Print Materials?

The inference that instructional designers use their own knowledge as a primary source when making revisions seems to be fairly robust. However, the analysis was taken a step further to examine the extent to which they did attend to the different data sources.

In order to answer the second question, revision segments were categorized according to the respective data sources.

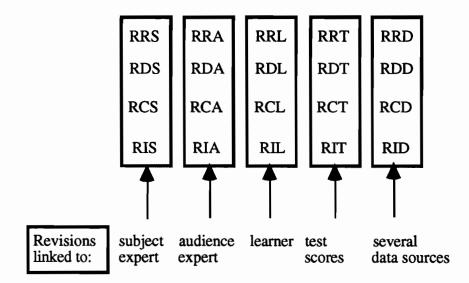


Figure 11. Revision codes linked to each source.

Revision statements were grouped according to the source (e.g., RRS plus RDS plus RCS plus RIS were all revisions linked to the subject-matter expert), shown by the boxes in Figure 11. Solution planning was again grouped with reviser's own input, since a check of the data found that it was always possible to link these segments to the reviser's own knowledge. Since revisers' input does not appear in the grouping to answer question 2, neither does solution planning.

The number of revisions linked to each external source was divided by the total number of revisions for each subject. This gave the percentage of revisions linked to each external source of feedback data. The results of this procedure are shown in Table 10 and represented graphically in Figure 12.

Table 10

Number and Percentage of Revisions Linked to Each Source of Feedback

		SUBJECT and TIME							
SOURCE		S1T1	S2T1	S3T2	S4T2	S5T 1	S6T1	S7T2	S8T2
Subject expert	(#)	1	0	4	2	13	8	6	14
	(%)	(0.6	0	1.6	1.8	10.2	6.3	4.6	10.3)
Audience expe	rt (#)	11	0.	29	21 -	9	12	4	18
	%)	(6.1	0	11.6	18.8	7.1	9.5	3.0	13.2)
Learner	(#)	32	1	21	16	16	3	25	31
	(%)	(17.8	0.8	8.4	14.3	12.6	2.4	18.9	22.8)
Test scores	(#)	2	0	3	4	4	0	4	3
	(%)	(1.1	0	1.2	3.6	3.2	0	3.0	2.2)
Several source	s (#)	3	0	2	7	13	0	9	6
	(%)	(1.7	0	0.8	6.3	10.22	0	6.8	4.4)

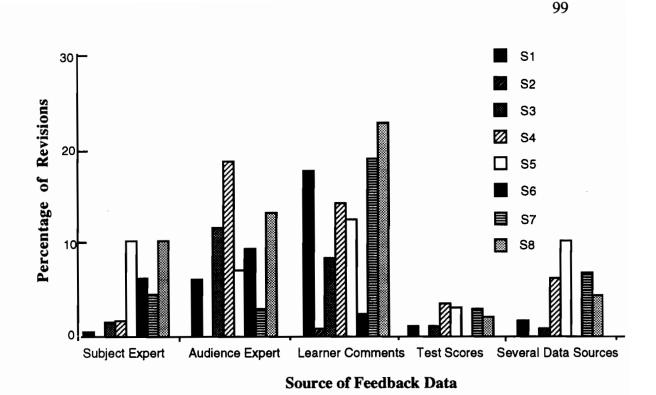
Another indication of the strength of the preferences is given by a calculation which is independent of the frequency of using external sources. When revisions linked to each external source of feedback data are expressed as a percentage of the total revisions linked to external data for each subject, the results shown in Table 11 are obtained. This shows that learner data, in the form of oral comments, were the first choice of five subjects and the second choice of two others, while the comments of the audience expert were the first choice as a data source for three subjects.

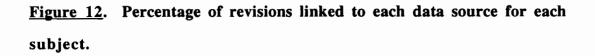
This finding of differences among the preferences of subjects is an example of a theme throughout the results of this study. Statistical analysis shows that the subjects were demographically similar, and allows them to be treated as a group. Closer examination of the protocols reveals individual differences, which will be identified as they emerge.

Table 11

<u>Revisions Linked to Each Source of Feedback Data as a Percentage of Total</u> of Revisions Linked to External Sources

SUBJECT and TIME							
S1T1	S2T 1	S3T2	S4T2	S5T1	S6T1	S7T2	S8T2
2.0	0	6.8	4	23.6	34.8	12.5	19.4
22.4	0	49.2	42	16.4	52.2	8.3	25.0
65.3	100	35.6	32	29.1	13.0	52.1	43.1
4.0	0	5.1	8	6.8	0	8.3	4.2
6.1	0	3.4	14	23.6	0	18.8	8.3
	2.0 22.4 65.3 4.0	2.0 0 22.4 0 65.3 100 4.0 0	S1T1 S2T1 S3T2 2.0 0 6.8 22.4 0 49.2 65.3 100 35.6 4.0 0 5.1	S1T1 S2T1 S3T2 S4T2 2.0 0 6.8 4 22.4 0 49.2 42 65.3 100 35.6 32 4.0 0 5.1 8	S1T1 S2T1 S3T2 S4T2 S5T1 2.0 0 6.8 4 23.6 22.4 0 49.2 42 16.4 65.3 100 35.6 32 29.1 4.0 0 5.1 8 6.8	S1T1 S2T1 S3T2 S4T2 S5T1 S6T1 2.0 0 6.8 4 23.6 34.8 22.4 0 49.2 42 16.4 52.2 65.3 100 35.6 32 29.1 13.0 4.0 0 5.1 8 6.8 0	S1T1 S2T1 S3T2 S4T2 S5T1 S6T1 S7T2 2.0 0 6.8 4 23.6 34.8 12.5 22.4 0 49.2 42 16.4 52.2 8.3 65.3 100 35.6 32 29.1 13.0 52.1 4.0 0 5.1 8 6.8 0 8.3





An analysis of variance was performed on the percentages of revisions from each source of feedback data to determine if there were any significant differences among the attention paid to each data source and to check for interaction effects of the time of providing feedback and for variations in the modules. The one way analysis of variance tests showed significant within-subject differences between the percentage of revisions linked to each data source in a number of cases. Of the ten tests performed, the six tests resulting in significant differences (p < 0.05) are shown in Appendices J through O. The differences are summarized in Table 12, which expresses revision statements linked to feedback data as a percentage of the total number of revision statements.

Table 12

Significant Differences in Revision Statements Linked to the Feedback Data Sources(p < 0.05)

Feedback source	Means (% of total revisions)	Probability of difference
Learner comments vs:	12.254	
Test scores	1.784	0.010
Learner comments vs:	12.254	
Subject-matter expert	4.415	0.028
Learner comments vs:	12.254	
More than one source	3.773	0.015
Audience expert vs:	8.664	
Test Scores	1.784	0.018
Audience expert vs:	8.664	
Subject-matter expert	4.415	0.045
Subject-matter expert vs:	4.415	
Test scores	1.784	0.040

The only effect between subjects was shown by the comparison between attention to subject-matter expert comments and to test scores, and this comparison also interacted with the module being revised. Inspection of the data showed that this effect was in the direction of Module 2. Further analysis showed that there was a weak interaction between Module 2 and the time of administering feedback. This is shown in Figure 13. These results indicate that subjects were significantly more likely to use subject-matter expert comments than test scores for Module 2 only, so that the main effect is spurious for this comparison.

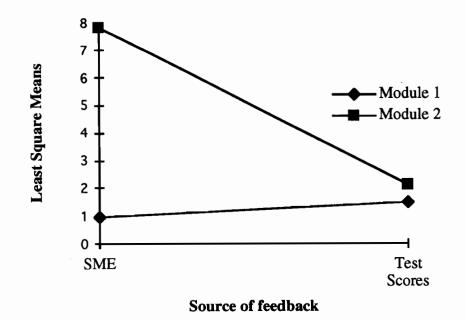


Figure 13. Effect of module on revisions based on subject-matter expert comments and on test scores.

The data analyzed for this question indicate that, although only about 21% of revisions could be linked to feedback data of any sort, as shown in Figure 10 and Table 9, when the data were aggregated across subjects, it appears that when they used feedback, revisers were most likely to attend to learner comments, followed by audience expert comments, followed by subject-matter expert comments, with least attention to test scores. Individual variation from this result has been noted in relation to Table 11. The data also show that there were significant differences within subjects' attention to five pairs of these sources.

Acceptance or Rejection of Feedback Data

So far, this question has examined the subjects' attention to each data source, without specifying what type of attention was paid. The total number of references to each data source was counted, and it was shown that, after their own input is discounted,

revisers paid most attention to learner comments and least to test scores. Closer examination of the think-aloud data to find out whether the attention reflected acceptance or rejection of the input from the data sources refines this picture slightly.

To investigate the attention more closely, revision statements were regrouped as shown in Figure 14. When revisions were initiated by the data source (RI-), or when revisers used feedback data to corroborate their own ideas (RC-), this showed that revisers were accepting the information from a particular source. When the revisers rejected feedback data (RR-) or when they made a revision despite feedback data (RD-), they were refusing to accept the information provided by the reviewer.

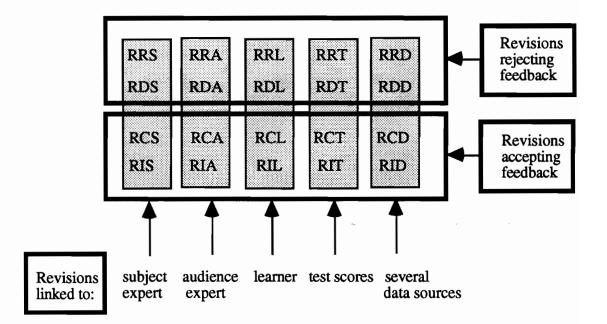


Figure 14. Revision codes reflecting reviser's acceptance or rejection of feedback data.

Revision codes were grouped in this way so as to be able to distinguish among the ways that subjects reacted to sources of feedback data. The total number of revisions referring to acceptance of feedback data from each of these categories is shown in Table 13

and those referring to the rejection of feedback data are given in Table 14, both grouped according to the source of the feedback data.

Table 13

CODE	S 1 T 1	S2T1	S3T2	S4T2	S5T1	S6T1	S7T2	S8T2	TOTAL
RIS+RCS	0	0	3	1	11	7	4	12	38
RIA+RCA	10	0	16	16	5	2	4	13	66
RIL+RCL	25	1	14	15	14	2	4	13	88
RIT+RCT	2	0	3	4	4	0	4	3	20
RID+RCD	3	0	2	7	13	0	9	6	40
TOTAL:	40	1	38	43	47	11	25	47	252

Revision Statements Initiated or Corroborated by Feedback Data

Table 13 shows that, again, subjects were more likely to regard learner comments positively, followed by audience expert input, then subject-matter expert input, then test scores, although there was a great deal of variation among the subjects. If the feedback sources are combined as "expert" (both subject-matter expert and audience expert) and "learner" (both oral comments and test scores) the difference in acceptance of feedback becomes less distinct. The total for acceptance of expert comments is 104 segments, versus a total of 108 segments for learner feedback. But revisers were given closely identified feedback data, rather than combined by expert or learner, so this more general grouping appears to be less relevant and will not be pursued further.

Table 14

CODE	S1T1	S2T 1	S3T2	S4T2	S5T 1	S6T 1	S7T2	S8T2	TOTAL
RDS+RRS	1	0	1	1	2	1	2	2	10
RDA+RRA	1	0	13	5	4	10	0	5	38
RDL+RRL	7	0	7	1	2	0	0	3	20
RDT+RRT	0	0	0	0	0	0	0	0	0
RDD+RRD	0	0	0	0	0	0	0	0	0
TOTAL:	9	0	21	7	8	11	2	10	68

Revision Statements Despite or Rejecting Feedback Data

Table 14 shows that subjects were most likely to reject data from audience experts, then from learner comments, then subject-matter experts, and that they rejected no data from test scores, nor when it came from more than one source.

These two sets of results show that, across the whole group, when subjects used feedback data they were more likely to accept feedback data than to reject them, although there is a wide variation in the range of acceptance or rejection. Subject S₆, used little external data and rejected more comments from the audience expert than she accepted. Examination of this subject's protocol explained this reaction. During the retrospective interview, the subject reported:

The --for some reason --the target population expert, I thought, wasn't very credible....'cause it's as if he wants to prove that he knows more than someone....it's almost like somebody asks you to critique it, so you go out of your way to find fault. I think he -- he wanted to earn his money....

A comparison of Tables 13 and 14 shows that in general, subjects were more likely to incorporate learner comments into their revisions and more likely to reject feedback from audience experts than inspection of the total attention to this source would suggest. It also shows that while test data were not used to any great extent, no test data at all were rejected. It also emerges from these two tables that subjects were more likely to accept data from more than one data source (RID + RCD) than they were to reject it (RDD + RRD) and that, in fact, they rejected no data which came from more than one source.

It should also be remembered that the results for question 1 showed that revisions from feedback data represent a small fraction of all the revision statements, and that the greatest number of revision statements come from the subjects' own input.

These data have reported the actual practice of the subjects during the revision task. To assess the relationship between subjects' perception of their behavior and their actual behavior, subjects were asked during the retrospective interview to estimate the time and attention that they paid to each of the sources of feedback data, and the priority they attached to each. Their estimates of the data they had used are given in Table 15.

Table 15

Subjects' Estimate of Their Attention to Each Source of Feedback Data

SUBJECT	ESTIMATE OF FEEDBACK USE	ACTUAL (from Table 13)
S 1	60 % of attention to learners	60% of attention to learners
S2	(subject did not use data)	
S 3	TPE and learner equally	TPE and learner equally
S4	TPE	TPE and learner equally
S5	SME, then learner, then TPE	learner, then several, then SME
S6	Learners, then SME, then TPE	SME, then TPE + learner
S7	Learner reaction based on test scores	several, then all equally
S8	Learners, then SME, then TPE	learner + SME+ TPE equally

Three subjects (S1, S3, S8) expressed their awareness of their use of learner feedback, without specifying whether that meant learner comments or test data. One (S4) overestimated his use of audience expert data and one (S8) overestimated her use of subject expert data. Only two subjects (S1 and S3) estimated their actual performance very closely. Question 4 will examine the reasons given by the subjects for the choices they made among data sources.

<u>Question 3</u>: Do Revisers Provided With External Feedback Incorporate As Many of Their Own Revision Suggestions as Revisers Provided With no External Feedback?

Two revision codes (RI + SP) occur whether or not subjects are given data. To answer the third question, the number of revisions proposed by the subjects when they were given feedback data was compared with the number when they were not given feedback data. Segments coded RI and those coded SP were grouped together for the "with feedback" condition for all revisers and compared with the same grouping for the "without feedback" condition. Table 16 gives the means for revisions based on the revisers own input with and without feedback data.

Table 16

<u>Comparison of Number of Revisions Linked to Revisers' Own Input in</u> <u>Both Feedback Conditions</u>

	With feedback data	Without feedback data
Mean	116.25	99.9
Standard deviation	47.49	35.58

A one-way analysis of variance showed no significant difference in the number of revisions based on the revisers' input whether or not they were given feedback (p < 0.05). In other words, they were equally likely to use their own input whether or not they were provided with data. For each subject, the feedback condition was provided for only one module and at only one of the sessions, that is, either for M₁ or M₂, and either for T₁ or T₂. It was therefore necessary to check whether the order of administering feedback or the module pairing had an effect on the number of revisions for each feedback treatment. The ANOVA showed that there was no effect caused by the time of administering feedback, nor by the module being revised, nor were there any interactions. These results are given in Appendix Q.

<u>Question 4</u>: How Do Revisers Establish Priorities Among Data Sources?

Three issues regarding reviser behavior enter this question, and three sets of data are available from revisers to consider it. These may be summarized as what they do, what they say, and what they say they do. More explicitly, there is first the record of what subjects do during the task set for them, provided by counting the revisions linked to each data source and already reported in the results for questions 1 and 2. Second there is the record of what they say during the the think-aloud in segments other than revision statements, giving reasons for the choices they are making. As far as can be determined, these responses are not edited. Third is their reflective account, made during the retrospective interview of what they say about their usual practice. The second and third data sets are necessarily more qualitative than the first one, and add information to the quantitative data of the subjects' practice already reported.

Use of Feedback Data in the Current Task

Since subjects were engaged in a think-aloud, they were not asked the reasons for their actions while they were performing the task. However, comments made during the think-aloud expressed the subjects' opinions of the data sources and the reasons for their choices among the feedback data. All relevant segments of the think-aloud will be reported verbatim to describe what the subjects said they were doing. The subject and session number are reported (e.g., S1T1) as well as the segment number (e.g., 1594).

Preference for Subjects' Own Input

Three subjects expressed a preference for using their own input:

S1T1		And I think I probably still would have had a tendency to avoid looking at that <i>(test)</i> data // well, I trust myself, I guess, enough to say that//
S2T1	3	but I won't look at other people's comments, //

	4	I prefer to improvise, OK?//
	693	Well, I felt more comfortable in a way, having no data, and having to rely on my own impressions. //
	694	Here, I feel like I have to put on put a brake on my own impressions, sort of, for a while, until I have dealt with this data and and incorporated it into, sort of, a strategy.//
	859	I'm lost by having too much data thrown at me. //
S7T2	486	Okay, now what I think I'm going to do is I'm going to skim through the text and just get a sense of my own personal my own "expert" comments. //

Yet it should be noted that S7 had earlier said:

S7T2	202	Well, it's much better having data.	(laughter)	//
	203	I feel much happier. //		

Several subjects expressed their awareness of their use of data to validate their own opinions and decisions:

S1T1	1110	it's sort of a validation thing here. //	
S7T2		Hey, that's what I thought too when I first read it! Vindicated. I feel vindicated.//	//

In spite of this expressed preference, only one subject described any familiarity with the content of the modules, and even so, there was an instance where his lack of content knowledge hindered his understanding of the structure of the module.

S1T1 956 And I'm not feeling on top of this enough to really say yet, whether -- whether I understand it all structurally. //

Most subjects represented themselves as "educated lay people" regarding the subject matter, and some even expressed a lack of any expertise.

S2T1	151	Okay, it's hard for me to correct because I am not familiar at all with this subject. //
	248	I have no level of entry at all here. Absolutely, no level of entry, //
S3T2	210	so I'm not an expert on that.// (referring to subject matter)
	1475	maybe it is the chemistry. //
	1476	I really don't have enough knowledge to know what it is.//
S4T1	563	and I have a sort of science background. //

S5T2	420	Well, I don't have the content knowledge to really know. //
S7T2	1864	Well, I mean, the other hang-up I have is chemistry. Right? (laughter) //

Opinions of the Expert Feedback Data

If subjects made any comment at all when they first read about the sources of

feedback data, they were generally positive comments:

S1T1	31	My first thought is this has been well tested //
	32	and that we can assume that the data here is pretty valid//
	33	and should be thought about when it comes to revising the material. //

Three subjects approved of the expert reviewers' qualifications.

S1T1 832	They both seem pretty well qualified.//
S5T1 15	the the subject-matter expert seems to have appropriate content knowledge; //
S7T2 210 215	I mean, I guess what the experts wanted is significant. // I must admit, already I'm prejudiced for the target population expert's comments, because they seem to have a training design background. //
217 218	because, although they get hung up on details, subject- matter experts will find flaws in the content. // So, we got to pay attention to them too. //
210	

The approval was often qualified. For example, later in the think-aloud, S₁ said:

S1T1 1113 I don't always agree with the target population expert, //

Other subjects were less impressed by the experts.

S2T1	21	So my subject-matter expert does not know my target population, //
	22	and my pseudo-expert on my target population mustn't
		know very much about my target population either.//

Referring to the audience expert, one subject repeated three times in different ways:

S3T2 316 Boy, this is a real hand-holder here, eh? //

Referring to the subject-matter expert, the same subject said:

S2T1	386	I'm looking for low scores again, //
S4T2	625	in terms of revisions, based on the students comments, I wouldI would definitely highlight anything that is like eighty percent and above, in terms of post test, //
S7T2	1952	See, my most important concern is that I've covered the test items. //

It should be noted that this was not supported by counting their revisions linked to the test scores.

Approach to Contradictory Data

There was no contradictory feedback data for subjects to deal with in Module 1: when more than one reviewer commented on the same passage, the information was generally consistent. In Module 2, experts and learners had reacted differently to an historical vignette (see Appendix B, p.10 for the text of Module 2). Learners said it was "confusing and irrelevant", while the target population expert said it was an "interesting historical approach." All four subjects in this condition dealt with this contradictory feedback, but did so in different ways.

S5 and S6 used their own judgement to supersede the conflicting comments.

S5T1	346	I think that's a really important point that might be stressed more, if you're dealing with non-science students.//
S6 T1	252	I think that thethe history section is probably the most interesting part of it. //
	253	And they don't like it because it's irrelevant. //
S7 chose to ac	ccept the	e learners' comments in preference to the audience expert's:
S7T2	1009	The target pop expert liked this,//
	1010	the learners don't.//
	1011	Get rid of it.//

S8 identified the contradiction and planned a compromise to accommodate both

learner and expert comments:

S8T2 117 My thought is basically to leave the sort of story format //

Table 17

Data Sources Usually Available to Subjects and Their Attention to Them

Subject	Data sources attended, in order of importance
S ₁	Learners, especially recent graduates, SME's
S2	SME's, other instructional designers, learner tryout, learner debriefing
S 3	Equal attention to SME and delivery expert
S4	Learners : SME's ; 2 : 1
S5	SME comments, small group learner tryouts
S 6	SME, instructional designers, text resources
S7	SME, instructional designers, learners (learners : technical; 4 : 1)
S 8	SME for content, learner comments for ID issues, test scores

There are obvious differences among subjects' reports of their usual practice, their account of what they were doing on this task, reported in the previous section, and their observed behavior, reported in Table 15. Three of the subjects (S₂, S₆, S₇) said that they usually consulted with other instructional designers, yet none of them linked this to the fact that they were instructional designers and this also constituted input to the task. In other words, "own input" was not reported as a source. An answer to the fourth research question: How do revisers establish priorities among data sources? demands more interpretation of the data than the answers to the earlier questions, and this interpretation will be treated in the next chapter. The interpretation will be based on the three sets of results identified so far: the actual performance observed in the results for question 2, the subjects' description of their actual performance, and their reports of their usual practice, both described in this section.

<u>Question 5</u>: When They Revise Materials, Do Instructional Designers Follow the Prescriptions of the Standard Model of Instructional Design?

Seven of the eight subjects had taken instructional design courses at the same University, all had used Dick and Carey's (1985) text at some time, and this text has already been identified as representing the standard model of instructional design. Therefore, in order to answer question 5, the subjects' actions were considered in relation to the instructional design model proposed by Dick and Carey, namely: (a) identifying an instructional goal; (b) conducting an instructional analysis; (c) identifying entry behaviors and characteristics; (d) writing performance objectives; (e) developing criterion-referenced test items; (f) developing an instructional strategy; (g) developing and selecting instruction; (h) designing and conducting the formative evaluation; (i) revising instruction; (j) conducting summative evaluation (Dick & Carey, 1985).

Of these, stages (e), (h) and (j) were not options for subjects in this study, because the formative evaluation had already been designed, and the revision aspect of the task was imposed on them. Summative evaluation was not possible, given the constraints of the research task. The actions and comments of the revisers will now be reported under each of the remaining headings, and summarized in Table 18. Then, step (i), revising instruction, will be analyzed more closely, using the textbook analysis shown in Figure 4.

Table 18 shows that there was considerable consistency among subjects regarding their observance of the instructional design model. All eight subjects were consistent in trying to identify the instructional goals (step a) in at least one of the sessions, either by asking the administrator for the overall goals, or by making assumptions as to what the goals were. None suggested conducting an instructional analysis (step b), and only one (S4) implied that she would check the entry skills of learners (step c). All were concerned with the specific performance objectives (step d) of the modules, although none related these to the two earlier stages of instructional analysis and entry behaviors.

Table 18

Step in ID model	S 1	S2	S 3	S4	S5	S 6	- S7	S 8
(a) identifying instructional goal	x	x	x	X	x	x	x	x
(b) conducting instructional analysis	-	-	-	-	-	-	-	-
(c) identifying entry behaviors	-	-	-	x	- .	-	-	-
(d) writing performance objectives	x	x	x	x	x	x	x	x
(f) developing instructional strategy	x	x	x	X	x	x	x	x
(g) developing, selecting instruction	x	x	x	x	x	x	X	x
(i) revising instruction	x	x	x	x	x	x	x	x

Revisers' Actions in Terms of the Standard Model of Instructional Design

All subjects commented on the lack of objectives or the need for objectives (step d) in at least one session, although several would have been prepared to accept objectives of a less rigid format than they might have been taught to write, given that this was an example of academic instructional material, rather than a piece of training material. Several subjects used the test questions so that they could infer what the objectives had been, and in some cases, this was the only reference to the test questions. One subject wanted to revise the test questions and was asked to concentrate on the module instead. Some subjects were so uncomfortable with the lack of objectives that they wrote their own. While the subjects were not charged with developing a brand new instructional strategy (step f), it is at this step that the instructional designer is most likely to be able to make revisions to text developed by a subject-matter expert, and this proved to be the case. All subjects suggested strategies to make the text easier to learn from. The option of developing new materials or selecting instruction from among preproduced materials (step g) was not applicable, since subjects were given a module to revise, rather than an instructional problem to solve. Media selection is a component of the development of instruction and instructional designers would be expected to offer opinions on this process. Subjects accepted the constraints of the print medium, but most made suggestions regarding a change in the format.

Thus the subjects in this study were reasonably consistent in whether or not they followed the steps of the standard model of instructional design. Whether or not they followed a common revision procedure is of more interest to this study. In Chapter 2, the case was made that most textbooks used in the training of instructional designers are weak in their revision procedures. Figure 4 summarized 16 textbooks published between 1971 and 1993, and the reader is referred back to this figure, which acts as a franework for considering the revision step (step i) of the standard model of instructional design.

Not all the steps listed in the textbook analysis are relevant to the results of this study. The first step (recognition of revision as a necessary component of instructional design) is irrelevant in that subjects were presented with a revision task, so there was tacit acknowledgement of the need for revision. Similarly, the second step (establishment of a theoretical base for revision) was also irrelevant, because subjects were asked to perform a practical task, and were asked not to analyze the process they were conducting. As practitioners, rather than theoreticians, they would not be expected to engage in developing a theory base, and use of think-aloud procedure mitigates against this. The third step (revision as a well-defined, well-articulated step) was established for the subjects: they were asked to enter the development process at the revision stage, not having been involved in the earlier stages. The last three steps in the textbook analysis (the presence of case studies or models, practice and feedback) are obviously part of the instructional strategy of the textbook, and are irrelevant in examining the actions of practicing designers. Their

presence or absence helps describe the extent to which revision is treated by the text, and therefore the effectiveness of the textbook's instruction.

All the textbooks investigated acknowledge the need for data collection: 13 specifying expert data and all 16 specifying learner data. The answers to the earlier questions show that the instructional designers in this study did use both these sources of data, but to a limited extent. In addition, "learner data" in instructional design textbooks specifies the use of test data. While subjects used test data as a way of establishing objectives, they did not use them as a source of many revisions. One subject asked if she could talk to the original author. She was told that she could propose this but that the author was unavailable, and she was willing to accept this.

The issue of providing a data collection instrument was not of importance, since feedback data had been collected for subjects, and the data collection conditions were given to them (see Appendix B, pp. 13, 14, 15). All subjects read this and accepted the procedures used. In addition, in the "without feedback" condition, once subjects had determined that this was all they would be given to work with, not one subject asked for data. Nine of the 16 textbooks propose a data handling procedure. For this study, it was decided to present the data as simply as possible, and this presentation was described in Chapter 3. None of the subjects asked for any more feedback data, nor for data organized in any particular way, such as the test item-by-objective summaries recommended by Dick and Carey (1985), nor was there any reference to the habitual use of such summaries in strategy statements (coded as ST) made by subjects.

The seventh step in the analysis was consideration of the question: Is a revision procedure given? Many texts which had given fairly detailed answers to earlier questions simply gave an instruction to revise, with no further advice regarding procedure. Nine texts gave some instruction (although this was often skimpy, at best), seven of these nine suggesting that the reviser plays some part in the process. The results of question 1 showed that the revisers in this study played a large part in the revision process, basing

almost 79% of the revisions on their own knowledge, and that whether or not they had received adequate instruction from textbooks in the revision process, they were confident revisers.

In summary, the instructional designers investigated in this study applied some of the standard prescriptions of the instructional design model and omitted others. Subjects were insistent upon identifying the instructional goal and learning objectives, and spent much time developing instructional strategies and reformatting the text to facilitate learning. On the other hand, they did not spend any time conducting an instructional analysis, nor did they propose item-by-objective summaries as part of the revision process. Further, they did not follow the textbook injunctions to use feedback data to the extent that might have been predicted.

Summary

This chapter has reported the data collected to answer the five research questions proposed for the study. The subjects in this study constituted a demographically similar group of professional instructional designers. Although the think-aloud technique was new for all of them, they were able to think aloud fluently. They understood the task assigned to them clearly enough that their definition of "revision" matched the definition envisioned for the research study, and they performed the revision task with confidence. The number of revisions ranged from 68 to 249 over the eight sessions, representing a mean of 32.5% of the theoretically useful segments across subjects.

Results for question 1 showed a significant difference between the subjects' use of their own knowledge and their use of external data sources when making revisions, (79% versus 21%, with a range of 58% to 99%, coming from their own knowledge). Analysis of variance for these data showed no significant difference among subjects. Analysis of variance performed for question 2 showed that, when subjects used external feedback data, there was a clear order of data use, with learner comments most used, followed by

comments from experts in the target audience, then subject-matter expert comments, comments from more than one source, and finally test scores. There were significant differences between pairs of these sources in five cases. Question 3 compared the number of revisions coming from the subjects' own knowledge when data were provided and when data were not provided. There was no significant difference between the number of revisions in each of these cases. Analyses of variance for each of these three questions also showed no effects caused by the two modules used, nor the time at which feedback was provided.

The think-aloud protocols provided a large body of data, which were obtained concurrently with the performance of the revision, and which were supplemented by comments made during a retrospective debriefing after each session. These data were used to answer question 4, which compared what subjects said about what they were doing with what they reported about their usual practice. Differences emerged among the three issues: what they did, what they said, and what they said they did. Question 5 used data from the think-aloud protocols to compare the subjects' performance with the standard model of instructional design already described in this study. This comparison showed consistency among subjects in following some of the relevant components of the model closely, while omitting others completely.

The final chapter will discuss these results and draw some conclusions about the performance of this cohort of instructional designers, with some generalizations about the actions of professional instructional designers.

CHAPTER 5: DISCUSSION

Overview of Study

The standard model of instructional design represented in most of the commonly used textbooks has a stage called formative evaluation which tells revisers to collect learner and expert data, with an emphasis on the pretest and posttest scores of learners, and to use these data to revise. However, not as much emphasis is given to revision as to the other stages of the instructional design process. This lack of textbook instruction on revision reflects the lack of research into revision. The literature gives evidence of the use of expert and learner feedback in making revisions, but demonstrates a lack of studies on the performance of practitioners, and no controlled studies to determine the impact of the revisers' own knowledge on their behavior. Consideration of this literature led to the general research question proposed for this study: What priorities are established among data sources when experienced instructional designers revise written materials?

While acknowledging the behavioral roots of instructional design, this study also drew from a broader theory base. Specifically, it applied the human problem-solving model and the think-aloud technique. The think-aloud methodology was chosen because it has proved useful in describing the solution of ill-defined problems, whether by novice or experienced performers.

Formative evaluation has been defined as a two-phase process; the collection of feedback data and its use in revision. The collection of feedback data from experts and learners was necessary before the research question could be addressed. This was described in Chapter 3 as "Phase 1." In "Phase 2", the focus of the study, eight demographically similar instructional designers revised two instructional modules, in one instance using only their experience of revision, and in the other, provided with feedback data from the sources recommended in a standard model of instructional design. All the data were supplied at the same time, so that revisers were forced to select among the

information they had been given. Verbal protocols collected while the subjects were thinking aloud were coded according to a problem-solving model, and revision statements were identified and tabulated. Descriptive and inferential statistical tests on the coded protocols allowed a picture of the performance of the revisers to emerge and this was amplified by examining more closely the data from the think-aloud and the retrospective interviews.

The results obtained in this study, and described in the previous chapter, will be discussed under the heading of each of the five research questions which guided the study. Since the problem-solving model was such an important foundation for the study, the chapter will continue by readdressing revision as a problem-solving act. The significance of the study, both theoretical and practical, will be addressed and some limitations outlined. The chapter will end with recommendations for further study.

Question 1: How Much Personal Knowledge Do Revisers Incorporate, if They Are Provided With Feedback Data From Experts and Learners?

The first question asked whether subjects used feedback data when such data are available to them. Results for the first question showed an overwhelming use by the subjects of their own input when they decided what revisions to make. On average, only 21% of revisions could be linked to the review data provided. Table 8 showed that this use of data ranged from a low of 1% for subject S2 to a high of 42% for subject S8. The analysis showed that there was no effect of "Time" nor "Module", nor any three-way interaction. The finding was consistent for both modules, and did not depend on whether feedback had been provided at the first or the second session.

It appears that even when they are provided with the feedback data identified in the instructional design literature as appropriate sources for informing the revision process, this group of revisers preferred to use their personal knowledge as a source of input for making

revisions. The think-aloud protocols show that subjects seemed unaware of the extent to which their use of their own experience was a factor in their revisions. It may be that revision procedures have been internalized to such an extent by these subjects that the procedures were being applied without being the topic of focal awareness, and therefore without appearing in the subjects' verbal protocols. It appears as if these subjects have reached Polanyi's level of "skilful performance", whose aim "... is achieved by the observance of a set of rules which are not known as such to the person following them" (Polanyi, 1958, p. 49). Further, no matter how little revision training they had received, these subjects had enough experience to have incorporated practical knowledge into the rules of their training. In Polanyi's words again: "Rules of art can be useful, but they do not determine the practice of an art; they are maxims, which can serve as a guide to an art only if they can be integrated into the practical knowledge of the art. They cannot replace this knowledge" (Polanyi, 1958, p.50).

Three subjects acknowledged the value they placed on their personal knowledge, as expressed when one subject said "well, I trust myself, I guess....." (S_1T_1 , segment 1596). However, no subjects acknowledged the use of their own input when they were asked during the retrospective interview what data they had used. It may be that they were unaware of their own input, or that these subjects use their own input so routinely that they do not consider it of note.

The finding that subjects used so much of their own input raises questions about the usual practice of revisers and about the nature of the data they rejected. These issues will be addressed later in this chapter, beginning with the attention that subjects paid to each data source.

<u>Question 2</u>: What Attention Will Instructional Designers Give to Each Data Source If They Are Provided With Verbal Comments From Subject-matter experts and Audience Experts, Verbal Data and Test Data From Learners, And Asked to Revise Print Materials?

Disregarding the subjects' own input as a source of revisions, and generalizing across the group, the data used by the revisers were, in order of use: the learner comments, followed by audience expert comments, subject-matter expert comments, comments from several sources, and finally, test scores. Moreover, there were significantly more revisions based on learner comments than on any other source except audience expert comments, and significantly less attention to test scores than to any other source except subject-matter expert comments.

In other words, more attention was paid to oral data, especially that from learners, than to achievement scores, and revisers did not use subject-matter expert comments to any extent. It is also notable that although subjects used pretest and posttest questions to define goals and objectives, the use of test data does not appear in the revisions to the extent that would have been predicted from the standard model. This is reminiscent of Dick's (1968) finding that revisers did not use posttest performance as a source of data, but since the injunction to use test data as a source of revisions is so prevalent in the training of instructional designers, it is a surprising result.

Several explanations suggest themselves for the lack of attention to test scores as sources of revision. It may be that the results of the posttests made revisers think that the draft instruction was adequate. For learning materials of this sort, a common standard is the attainment of 80% on each question, and one subject (S4) specifically mentioned this criterion during the revision. Examination of Table 2 shows that for Module 1, the success rate was greater than 80% for 12 of the 23 questions, and for Module 2, the success rate was greater than 80% for 6 out of 23 questions. Further, on the matched pretest/posttest questions, scores decreased or stayed the same for three items in Module 1, and for two

items in Module 2. These observations suggest that the original modules were susceptible to improvement.

Another explanation might be that the revisers did not value the quality of the tests, and indeed, it has already been noted that one subject asked to be allowed to revise the tests. Yet subjects used the test items extensively to help them decide on the objectives of the modules, and protocols reveal segments (coded as VST) where subjects were reading test items, or the scores related to them. The lack of corresponding segments which could be coded as RIT (revisions initiated by test scores) suggests that although subjects were reading the test items and scores and using them to identify objectives, they were not explicitly basing revisions on test scores in the manner that would have been predicted from their training.

The analysis of variance for question 2 showed the only effect of the two different modules on the revisions made. Figure 13 showed that subjects made less use of subject-matter expert data for Module 1 than for Module 2, no matter which session Module 1 was given to them. Examination of verbatim statements related to subject-matter expert comments in the transcripts (coded as VSS), showed that subjects given feedback on Module 2 read more subject-matter expert comments than those given feedback on Module 1, although the possible number of comments was almost identical for each (39 for Module 1, and 41 for Module 2).

Only two (S3 and S4) of the four subjects given feedback on Module 1 commented directly on the qualifications of the subject-matter expert who had reviewed the module. As already quoted in Chapter 4, S3 was critical of the subject-matter expert's qualifications (a B Sc in Nutrition), referred to her as "our little subject-matter expert" (S3T2, segment 928), and did not find her comments "particularly astute" (S3T2, segment 985). Subject S4 simply said "I don't find the SME's comments useful" (S4T2, segment 582). The qualifications of the expert reviewers are given in Appendix B, and Sheet 13 (M1) and

Sheet 14 (M1) show that the qualifications of the subject-matter experts for the two modules are almost identical.

A major difference between the subject-matter expert comments for Modules 1 and 2 was the quality and usefulness of their comments. The subject-matter expert for Module 1 did not make many revision suggestions, as was noted in the description of Figure 6, and many of her other comments were generally supportive of the content; many comments just said "True" or otherwise endorsed the information in the module,. The subject-matter expert providing feedback on module 2 contributed more suggestions for revision, and these seem to have been accepted more readily by the subjects.

Table 11 showed that the audience experts' comments were the first choice of feedback data for three subjects and the second choice for two subjects, although a comparison of Tables 13 and 14 shows that revisers were more likely to reject data from audience experts than from any other source. Only one subject (S₆) expressed any criticism of the audience expert, and she rejected as many suggestions from this source as she accepted.

The expert reviewers had been carefully matched regarding their qualifications, so it seems as if the revisers' perceptions of the qualifications of the experts, and of the worth of their comments were more important than the actual qualifications when it came to accepting or rejecting their comments. These issues will be readdressed when the implications of these findings are discussed.

As well as considering what feedback sources are attended by the subjects, it was interesting to ask whether this attention was acceptance or rejection. Comparison of Tables 13 and 14 showed that, overall, subjects were generally more likely to accept data from any particular source than to reject it. From Table 13, there was a total of 252 revisions initiated by feedback (coded RI-) and corroborated by feedback (coded RC-). Table 14 shows that there were 68 revisions made despite feedback (coded RD-) or rejecting feedback (coded RR-). It is apparent that when subjects attend to data they are more likely

to use them than to reject them, and that they use data to initiate revisions or to corroborate their already formed opinions.

It is worth noting that 68 revisions linked to external sources of feedback (about 21%) were made despite the feedback. This further supports the observation of the subjects' use of their own knowledge. Here it was being used to override the information given to them, and might be regarded as an aspect of the subjects' response to contradictory data.

Throughout this study, results have been reported for revisions based on data from more than one source, coded as RID, RCD. The relatively large number of these revisions, compared with those from other sources (see, for example, Table 13), suggests that subjects were likely to assign value to suggestions coming from more than one feedback source, and to incorporate a revision if the suggestion came from a number of review comments (often including their own comments). This supports Kandaswamy's (undated) assertion that revisers perform a meta-evaluation of the feedback data they are given. It also adds to the importance of the reviser's own input. Not only were the subjects making a large proportion of the revisions directly from their own knowledge, but also this knowledge was mediating the revisions based on the feedback provided.

Question 3: Do Revisers Provided With External Feedback Incorporate As Many of Their Own Revision Suggestions as Revisers Provided With no External Feedback?

The first two questions used data collected only in the "with feedback" condition. For the third question, a comparison was made between the revisions linked to revisers' own input when they were provided with data and those made when they were not provided with data. The question grew out of the predictions that revisers would use more of their own input when they were not given data to work with, and that there might be a

learning effect during the first session, so that they might make more revisions based on their own input during the second session.

The average number of revisions in each case was reported in Table 16 and the analysis of variance in Appendix Q. This analysis showed no significant difference among the number of revisions in either case, and no interactions caused by the module, nor by the time of administration of the data. Subjects in this study were equally likely to use their own knowledge whether or not they were given data. This suggests the conclusion that there is no reason to collect review data to give experienced revisers, but this might be too hasty a conclusion. To note only one reason at this point, the quality of the revisions has not been considered.

Subjects did not use their own input exclusively, with the exception of one subject, S2, who made all revisions before reading the feedback data, and then made only one change based on data. It is clear from the results for question 2 that revisers <u>do</u> use data, if only for, on average, about one fifth of the revisions that they make, although one subject based as many as 42% of her revisions on feedback.

Revisers have been quoted as saying that they felt happier with data, that the comments "vindicated" or "validated" what they already thought. The only negative comment in the "with data" condition came from the subject who said: "I'm lost by having too much data thrown at me" (S4T₂, segment 859). Provision of all the data at the same time was deliberate, to force subjects to choose among the information they were given.

Question 4: How Do Revisers Establish Priorities Among Data Sources?

So far, this discussion has examined the feedback data sources used, the extent to which they were used and the purposes for which they were used. The fourth question investigated the decisions made by subjects regarding the incorporation or rejection of various types of data The first thing to emerge was evidence of contradictions among subjects' perceptions of what they had done in this task, their practice as evidenced in their verbal protocols, and their descriptions of their usual practice, an observation also noted by Rowland (1992). Results for question 1 showed that, in fact, the revisers' own knowledge was the predominant source of revision statements. Results shown in Tables 13 and 14 suggest that they were more likely to accept than to reject the data they were given. Even so, the predominance of revisions linked to the subjects' own input shows that they appear to value their own resources most highly as a data source.

Verbatim statements from the think-alouds and from the retrospective interview have quoted the subjects' perceptions of what they were doing in the task, and their reports of their usual procedures, and examination of these statements in relation to the actual performance gives some evidence for the priorities established. From the protocols, it emerges that subjects were more likely to be sympathetic to learner comments. Reasons given by revisers for their use of various sources of feedback data in this study were based on the value they placed on each of the data sources, and did not closely match their reports of their usual practice. Neither did they report that they had used their own knowledge as input into revision decisions.

Use of Expert Feedback

Subjects reported that they had little knowledge of the content of the modules they were revising, which suggested that they might rely on subject matter comments. Also, when they were asked what feedback data sources they usually had available and which ones they were most likely to use, Table 17 shows that the majority said they used subject-matter expert data as a source of revision suggestions. In Table 15, subjects were shown to be reporting a heavy emphasis on the use of subject-matter expert input into revisions. Yet the actual revisions based on subject-matter expert review comments was only 4% of

the total revisions, significantly less than those based on learner comments or audience expert comments or their own input.

They also reported their usual use of other instructional designers during revision, although they were given no feedback from instructional designers in this study so as not to contaminate the study with input from the practitioners in the same domain. Their reporting of input from other instructional designers may result from the team approach reported by Nathenson and Henderson (1980) which is common in the large companies in which most of the subjects were employed. It may also be an oblique acknowledgment of their own input into the revisions.

Use of Learner Feedback: Oral Data

When subjects used the data they were given, they were most likely to use the comments that had been made by learners. Table 17 shows that three subjects (S₁, S₄ and S₇) said they usually paid most attention to learner data, although comparison with Table 15 shows that only one of these three (S₇) actually did so in the research task, and a different three (S₁, S₅ and S₈) estimated that they had paid most attention to the learner comments in the research task. Two subjects explicitly said that they were using a strategy which involved acting like learners, both in conditions where they were not given feedback (S₄T₁ and S₅T₂).

When subjects expressed an opinion about learner comments, the opinions were generally positive. One subject explained his willingness to accept learner comments by saying: "these general comments, I think, are --are maybe all the more strong, in the sense that, often, students will just expect stuff to be like this" (S4T2, segment 481). Yet Table 15 shows that this subject estimated that he had given more attention to the audience expert in the research task, and that this estimate was correct.

Use of Learner Feedback: Test Data

The subjects used far less test data to plan revisions than might have been predicted from the attention given in textbooks to gathering test data during formative evaluation. The case has been made for allowing the subjects to define their own task, so no attempt was made to define objectives for them. In the absence of defined objectives, subjects read the test questions to help them identify the objectives of the modules. Having identified the goal of the module as changing learners into better consumers, they criticized the test questions for being too factual, which may have caused them to place less importance on the information given by the test scores.

However, two pieces of data suggest that this was not an anomalous observation. When they were asked to report their usual practice, Table 17 showed that only one subject (S8) made clear reference to the use of test scores. All other reference to the use of learner data was ambiguous as to whether test data or learner comments were being used. This may mean that instructional designers do not place as much emphasis on test data as the standard model of instructional design suggests. Also, it was reported in Chapter 4 that no reference was made to item-by-objective summaries of test questions, either in the task at hand, or in their usual practice. This presents a picture of an instructional designer who does not follow the standard model of instructional design. This issue will be readdressed in the discussion of the fifth question.

Use of Contradictory Data

There was only one example of contradictory data resulting from the review, so it is difficult to answer the question of how subjects reconciled such a problem. In the particular case described in Chapter 4, it was clear that the four subjects confronted with contradictory data from learners and the audience expert were not disturbed by it. One sided with the learners, two accepted the audience expert's comments, apparently because these comments supported their own beliefs, and one planned a compromise. When

feedback data contradicted the subjects' own opinions, Table 14 shows that revisers used their own knowledge in preference to the external data. It should be noted that the review data for this study were real, not contrived, and contained consistent comments, so more research with inconsistent data is necessary before further conclusions can be drawn.

In summary, three contradictory factors seem to be present in the results for question 4. First, subjects generally lacked content knowledge in the materials being revised, and were slightly more familiar with the audience for the modules. This suggested that they might rely more heavily on the subject-matter expert's feedback than on the audience expert's, while, in fact, they used feedback from the audience expert in preference to that from the subject-matter expert. Second, the subjects in this study reported heavy use of subject-matter expert input in their usual practice, although this was not supported by the results of this study. Third, instructional designers are instructed to value test data from learners when they revise, and these subjects made minimal use of test data for making revisions.

Gilbert's (1992) report of "a small but sure negative correlation between how exemplary performers actually do their jobs and how they say they do them" (p. xv), and Duy's (1989) observation that revisers do not accurately estimate the extent to which they use a data source are supported to some extent by the results of this study. None of the revisers reported their own input as a data source during debriefing, although most of the revisions are linked to their own input. But after this important consideration is removed, Table 15 showed that only two of the subjects assessed correctly that they had placed most emphasis on learner data during this task, although they did not specify whether these data were test scores or learner comments.

Subjects' principal data source was their own knowledge, even when they were given data. When they chose among sources of feedback data, they were willing to give more credibility to the learners' comments, were generally unimpressed by the experts'

comments and credentials, and therefore less likely to accept their input, and did not use the test scores to any extent, except as will be explained in the discussion of question 5.

<u>Question 5</u>: When They Revise Materials, Do Instructional Designers Follow the Prescriptions of the Standard Model of Instructional Design?

All subjects reported that they had used one textbook, (Dick & Carey, 1985) at some time, either as a required text during their training, or as a reference book since then, which substantiated the use of this text as the standard model of instructional design used in this study.

Instructional designers are trained to use subject-matter experts as data sources, and these subjects reported that they used such data routinely, yet their behavior in this task does not support this. Also, when instructional design textbooks address the issue of revision, all recommend the use of pretest and posttest scores as an estimate of how successfully the materials teach. Table 18 summarized the subjects' attention to the standard model, and showed that it was attended by all subjects to the same extent. That is, all subjects followed all but two steps in the model: conducting an instructional analysis and identifying entry behaviors.

It may be that, knowing that they were dealing with undergraduates, subjects assumed that all the learners would be a homogeneous cohort, and that there was no need to identify the entry behaviors, although one subject mentioned that this might be necessary. Performing an instructional analysis is an early stage in the systematic design of instruction, and one which follows logically from the definition of objectives. No mention was made of this step by any of the subjects, although every one of them talked at length about the need for objectives. It is possible that instructional analysis was omitted because it is an early stage in instructional design; the subjects may have perceived themselves as coming into the project at the end, so that they did not see the early steps as being part of their task. It is also possible that an instructional analysis is too demanding of theory to form a regular part of the activity of an instructional designer and that this is one of the corners that has been cut in practice. Finally, subjects may have assumed that the instructional analysis had already been done, although no-one checked this fact with the administrators.

The revision step of the standard model gives another example of subjects' divergence from their training. Results have shown that they did not use review data, especially test data, to the extent that they are told to in the model. The lack of objectives made a test item-by-objective summary impossible, but none of the subjects mentioned this as a problem for them, nor referred to it in any way. Again, the fact that they were coming into the project at the end of the process may have affected the subjects' use of data. It has been noted that several subjects gave themselves the role of learner in their first approach to the materials. Their need for ownership of the modules may have affected their attention to the various data sources.

These results suggest that the subjects followed the model, but in a much looser way than the systematic approach suggests. They showed a preference for "soft" oral data rather than "hard" test data, in a similar manner to that reported by Ciesla (1976). They were very organized in approaching the revision, but did not use such organizing principles as an instructional analysis or detailed analysis of test results. This seems to provide another example of Shanteau's (1992) observation of the use of heuristics, rather than algorithms, by experienced performers faced with ill-defined problems. It is typical of practitioners giving reasons for their actions that "lore and experience are cited more frequently than are hard data" (Geis, 1986, p. 4). It is also typical of experienced practitioners to have internalized rules to such an extent that they would apply them unconsciously. Duchastel has noted that professionals have a strong cognitive need for "a sense of control over the analytical process involved in problem-solving" (Duchastel, 1990, p. 439). It appears as if these instructional designers were experienced enough to be able

to take short cuts in their work, and to have distilled from their training the principles they perceived to be of most use to them, giving them the control that they needed.

In developing an instructional design model to teach problem solving, Dijkstra wrote: "Sometimes within a community of problem solvers there is agreement about the problem-solving procedure, sometimes the problem solver is free to develop a method to solve the problem" (Dijkstra, 1991, p.6). In the absence of prescriptive literature on revision, it would appear that revisers are free to develop their own methods of problem resolution. How far do the revisers studied constitute a community with agreement regarding procedures?

Subjects were a demographically similar group of experienced instructional designers who had received similar training. Subjects' reports of their training supported Ely's (1992) opinion of the preparation of educational technologists. Ely wrote: "The concepts and procedures [of evaluation] are incorporated into other courses, but they tend to have a minor place in the entire professional education curriculum." (p. 27). Subjects reported use of a similar list of textbooks, but no one mentioned Gropper (1975), the only book dedicated to revision. The selection criteria for subjects--all were practicing instructional designers who belonged to the local chapter of NSPI--explained why they belonged to similar organizations and read similar journals. Their use of journals and interest in conference attendance demonstrated a degree of professionalism common to all of them, and they were taken to constitute a community of instructional design

According to Alexander, Shallert, and Hare (1991), members of a community share a body of tacit knowledge that they call "sociocultural knowledge" (p. 325). In relation to composition, Flower, Carey and Hayes (1985) express this type of knowledge in a different way when they write about "personal maxims", identified by strategy statements (ST) in this coding scheme. Comments made by the subjects suggest a common view of the task, of the requirements for revision, and of the means to effecting the necessary

revisions. They used the data, especially data from several sources, and data which supported their own views. No examples could be found where revisions were made despite the reviser's opinion, so the code for this does not appear in the results. They used their experience, and the protocols are replete with strategy statements (ST) and knowledge statements (KS) describing this experience. Chapter 4 has described how they proposed revisions based on learning principles such as the need for better formatting, presentation, feedback and practice, and active learner engagement.

The results of this study seem robust enough to call the subjects a community of instructional designers fitting Winn's description of "successful professionals" who solve problems by adding "experience that is often tacit or intuitive" to "a mastery of theoretical principles" (Winn, 1990, p. 66). They generally followed the same model of instructional design. They were experienced enough to have modified the prescriptions of their training, and had done so in a fairly consistent way.

Revision as Problem Solving, and the Use of the Think-aloud Technique

In Chapter 2, revision was characterized as an ill-defined problem, and the subjects' solution of the revision problem was typical of the process for solving ill-defined problems. The representation of the task was modified as the solution proceeded, as observed by Simon and Hayes (1976), although once established, the definition was stable across the duration of each session, implying that the subjects had a clear picture of the task. Gap-filling decisions were made (Hayes, 1989) as subjects satisficed on a solution. Different solutions were proposed to the revision problem (Simon, 1973), although a description of these solutions is outside the scope of this study.

Problem-solvers build their own representation of an ill-structured problem, both at the beginning of the process, and as the solution proceeds (Geis, 1986; Simon & Hayes, 1976) and this behavior was observable among these subjects. There was some alteration in task definition as the revision proceeded, consistent with Simon's (1978) observation that "As recognition of particular features in the situation evokes new elements from long term memory, the solver's problem space undergoes gradual and steady alteration (p. 287)." Part of this alteration in problem space serves to define the problem more clearly, and part of it helps to structure the problem for the solver and to identify operators to search the problem space. It was noticed that some subjects' protocols contained segments related to task clarification (TC) after the task had proceded for a few minutes, as well as task representation segments (TR) all the way through. In this way, subjects were refining their task definition and redefining the problem space they were using.

Characteristic of the behavior of experienced performers, rather than novices, subjects generally spent considerable time setting up the problem space (Simon, 1973), as reflected in the number of segments coded as task representation (TR) and task talk (TT) early in the protocols. They were also experienced enough in the task to be able to call up previous problem spaces and operators from long term memory, indicated by the facility with which they proposed some revisions (use bullet points, shorten the paragraphs, etc.,) and by the strategy statements (coded as ST) in their protocols ("I usually...", "At this point I would....", etc.,). In Duchastel's (1990) terms, they were re-exposing themselves to their previous experience, and in Rowland's (1992) terms, they were retrieving a template to assist the solution.

Subjects' casual use of the terms review, revise and edit reflected the lack of precision with which these terms are used, and perhaps also the demands on an instructional designer, who is expected to perform all these functions while evaluating materials. It was possible to clarify the use of these terms--and therefore the research task-for each subject without constraining the task too closely. After this clarification, all subjects were able to define the task clearly for themselves as a revision task, and all used basically the same definition of revision as that used for the study.

The coding scheme used for the transcribed protocols was based on the premise that revision as a component of formative evaluation was an example of problem solving. The stages of problem solving outlined in the coding scheme were clearly identifiable in the transcribed protocols. There were instances when revision statements were not preceded by problem statements (i.e., problem resolution was occurring without overt problem identification), a feature of protocols obtained from experienced performers that has also been noted by Flower and Hayes (1980). The finding of parallels between the protocols obtained from this study and those reported by Flower, Hayes, and their associates lends justification to the use of the same problem-solving model as that used in composition.

Data collection followed Ericsson and Simon's (1980; 1993) advice to use thinkaloud procedures to capture the procedures of experienced performers. In general, subjects proved to be good at the think-aloud task, although it was a new experience for all of them. The "with feedback" condition seemed to be more demanding of subjects, presumably because more pieces of information had to be kept in focal awareness. The mean time for this condition was longer, and subjects needed more prompting to think aloud when feedback data were available to them.

Even though every attempt had been made to organize the data so that subjects were not overloaded, they still had to deal with a number of different pieces of paper when they were given feedback data, and there was more tendency for verbal production to shut down in the presence of all the material necessary to provide the feedback. Subjects used a variety of strategies for dealing with the different pieces of feedback data. These included physically moving papers around, taking apart the stapled papers and spreading them out so as to be able to see as many as possible, reading each set of comments sequentially, or reading each set of comments on a particular section before going on to the next section. It appeared that these strategies were effective, and that the process of thinking aloud was not adversely affected by the provision of feedback, since the greatest number of prompts for any subject never exceeded 18 in 2.5 h, or one every 10 minutes. The median value was about one prompt every 50 minutes.

Subjects were told that they could make notes while they were revising, as long as they continued to think aloud. Examination of these written notes did not provide information on the process that they had followed, and in many instances, gave only a sketchy picture of the product. Information provided by subjects during the retrospective interview was often inconsistent with the evidence from the think-aloud protocol, substantiating the view, already expressed, that experienced performers are often poor reporters of their actions. Thus the think-aloud procedure, although it is probably not providing a complete picture of the subjects' cognitive processes, appears to a reliable source of information on the complex task of revision.

Summary

The general question which guided this research was: What priorities are established among data sources when experienced instructional designers revise written materials? The research began with an assumption from the literature that instructional designers who had been trained according to a standard model would do what they had been trained to do; they would use external review data provided for them and incorporate it into revisions. If they were not given data, they would act as their own reviewers, and there would be evidence of differences between these two conditions if their actions were tracked by asking them to think aloud.

The results of this study have shown that the original premise was only partly correct. Instructional designers do follow certain procedures that they were trained to follow, but not to the extent that had been anticipated. The instructional designers in this study used feedback from experts and learners to a limited extent, and placed far more reliance on their own input. They acted as their own reviewers whether or not they were given data, sometimes acting the role of learner, and there was little difference between their performance in these two cases. Moreover, they were not aware of the discrepancies between their actions and their descriptions of their actions.

At first sight, it appeared as if the subjects in this study were using their own knowledge to the exclusion of nearly all the feedback data they had been given when they revised the instructional materials. Closer investigation shows that although they gave priority to their own knowledge as a source of revisions, they did, in fact, use the sources of feedback recommended in the standard models of instructional design. However, they may not have used these data for the purposes usually recommended (e.g., test scores were used to define objectives, not to make revisions), nor to the extent that the standard models would suggest. They sometimes used feedback as validation of their own ideas, and felt more comfortable when data were provided. They were experienced enough to have internalized the maxims of their training and to be able to apply these maxims confidently and unconsciously.

Specific Contributions of this Study

The findings of this study have implications for both the theory and practice of formative evaluation, especially the revision phase of formative evaluation, and will be outlined under the headings of theory and practice.

Theoretical Significance

This study fits into the programmatic research of a team which has as its long term goals the identification of the effectiveness of different formative evaluation strategies, and the establishment of guidelines to assist revisers of instructional materials. The team has been in existence for ten years and has so far investigated the collection of data from experts and learners, according to the prescriptions of the systematic design of instruction. Appendix A gives an overview of the work of the team.

The present study has extended this programmatic research from a focus on the collection of feedback data to the use of these data in revision. It has continued to separate data collection from data use in formative evaluation so that each of these steps may be

examined more closely. It has also drawn from the theory bases of problem solving and composition and, therefore, from their methodologies, specifically the use of the thinkaloud procedure. Its particular contribution to the research program has been to examine the actual practice of experienced instructional designers.

The review of literature on revision as a step in formative evaluation suggested a gap in the research on the process used by trained revisers when they revise print materials. The research questions in the present study investigated the information attended by experienced instructional designers when they revise printed instructional materials. Unlike most previous research into revision, this study gave professional instructional designers a naturalistic revision task, where they set their own task definition, and were not constrained to use or to avoid particular data sources. It demonstrated that experienced revisers incorporate their own knowledge into revisions, rather than using suggestions made by reviewers, and raised questions about the type and quality of data collected for revisers' use.

Clear evidence of the stages of the human problem-solving model in the protocols of these subjects substantiates the use of this model as a theoretical framework when studying instructional design and specifically formative evaluation. Since model building and heuristic development are stages in the process of theory building, and since little work has been done to test the heuristics applied by revisers, it is not surprising that the act of revision in formative evaluation has so far had a limited theoretical base. Placing it within a problem-solving paradigm and relating it to composition has allowed revision to be explored using the methodologies common to these two domains.

The human problem-solving model has also enabled the development of a more finely grained coding system than in earlier studies. For the first time, the revision behavior of instructional designers has been tracked in great detail. Instead of accepting revisers' reports of the feedback data they had used, this coding scheme allowed revisions to be linked to specific sources of feedback. The use of problem solving as a model

allowed a robust and fine-grained coding scheme based on a firm theoretical base. Thus it provides the first empirical research into the actions and preferences of experienced revisers during formative evaluation under conditions as naturalistic as possible. In this way, it has extended the study of expert performance into instructional design, adding a new data set to the study of expertise in different domains.

Much of the research into revision has been focused on the behavior of writers, whether novice or expert, while they are revising their own original compositions (Fitzgerald, 1987, summarized revision research to that date). But other professionals engage in revising, and instructional designers are often charged with revising the writing of others. This study provided a link between an existing body of research and a less examined area, that of "adults in occupations where functional writing is prevalent" (Fitzgerald, 1987, p. 498).

Practical Significance

The results of this study also offer implications for practice. The most salient result, of special interest to anyone charged with managing the design of instruction, was the extent to which revisers use their own knowledge base when they revise. Across the eight subjects, nearly 80% of the revisions were based on the revisers' own knowledge. Since the revisers' input is so pervasive, this suggests that a project manager should choose the reviser carefully.

Yet reviser knowledge is not the only source of the revisions that were proposed. Figure 4 showed that, without exception, textbooks advise instructional designers to collect and use feedback data to drive revision. Because collecting revision data is an expensive and time-consuming process, it is often omitted in materials development. The results of this study suggest that data collection is worthwhile, because although external data were incorporated in only about 20% of the revisions, revisers said that they felt more comfortable when they had data. It is also worthwhile to collect data from a variety of sources, since it appears that revisers value data if it comes from more than one source. The study leaves open the question of whether such feedback should be collected by the classic methods outlined in instructional design textbooks, or by newer methods, such as those summarized by Tessmer (1994).

It is also clear that expert reviewers should be carefully chosen. The credentials of the reviewers were important to the subjects. Their assessment of the reviewers' credentials and the value of the comments made by the reviewers affected the attention subjects paid to the review comments, whether or not the review data were instrinsically valuable.

The practice of training instructional designers is well established and this study has shown that, in general, instructional designers follow the model on which their training was based, although they omit some of the steps, either consciously or unconsciously. This result points out a question of practical importance in the training of instructional designers. If evaluation of the revisions show that omitting these steps does not have any effect on the revised product, it may be possible to streamline the model that is taught to neophytes. If the revised materials show that the steps are, in fact, crucial, it is necessary to reconsider the way they are addressed in training.

Limitations

Some of the limitations of the study and the actions taken to control these limitations have been described in the relevant sections of the text. For example, the small number of subjects was accommodated by counterbalancing the treatment conditions, which matched subjects with themselves within treatments, giving 16 data sets for 8 subjects. In addition, use of both the think-aloud procedure and a retrospective interview produced a large amount of rich oral data from a small number of subjects.

While the think-aloud procedure has been criticized for distorting production, there is enough evidence to warrant the use of the technique to capture processes that would be

invisible in the final product (e.g., Swarts, Flower, & Hayes, 1984), and the results of this study support this use. The think-aloud was supplemented by the subjects' written notes, and by the notes taken by both researchers. The written trace of the revisions produced was not helpful in identifying the processes followed by the revisers, but was used to help in transcribing and coding the oral data and provided helpful clues regarding emphasis and context.

The collection of feedback data by different researchers may be seen as a limitation to the validity of the data, but strict attention was paid to using matched procedures for each of the studies to ensure that data sets were comparable for the two modules.

Formative evaluation is traditionally considered to be an iterative process (Geis, 1986), and models of instructional design advocate collecting data and using them for revision at each stage of materials development. In this study, revisers were deliberately provided with all data at once to force the revisers to be selective and to prioritize.

This study's place in a program of research could be identified as a limitation, since each researcher deals with only a small section of the whole problem, but it also constitutes a strength. It has enabled the researcher to use data sets collected for earlier projects (Israeloff, 1992; Rahilly, 1991; Tremblay, in progress) and to build a coding scheme based on earlier productive schemes (Duy, 1990; Rahilly, 1991; Saroyan, 1989). It has also provided a pool of fellow researchers for assistance and for formative evaluation of the work in progress.

Recommendations for Further Research

This study has collected an extensive data base which is now susceptible to further analysis. For example, it is recommended that the segments shown in Appendix E which are not coded as revision be examined more closely. Protocols could be re-examined in detail for the problem-solving strategies employed by the revisers. The protocols could also be examined for evidence of expert performance compared with the standard definition of expertise in works such as Chi, Glaser and Farr (1988) to develop a model of expertise in instructional design, or at least in revision as a component of instructional design. According to this model, the "most expert" and the "least expert" revisers could be identified from the eight subjects, and the revisions they proposed could be tested on learners. It would also be possible to develop a model of the knowledge base of professional instructional designers by investigating all the knowledge statements in the protocols.

From analyses such as these, a revision heuristic could be proposed which would be a valuable step in planning the training of neophyte revisers. A model of reviser behavior would include the most effective characteristics of a good reviser. Combined with an evaluation of the revisions produced, this would help in the selection of revisers, which has been shown to be an important consideration.

Throughout the study, a case has been made for considering the subjects as a demographically similar group, and their similar approach to the revision process was noted in Chapter 4. However, individual differences were apparent, and have been noted where they occurred. The differences offer intriguing insights into the individual reviser's performance. For example, subjects S₂ and S₈ provide an interesting contrast in style and performance. S₂ based 99% of revisions on her own input. She had the most irrelevant talk during the think-alouds: adding the number of segments coded as dialog (D) with unrelated talk (UT) over the two sessions gave a total of 274 segments. She also had the most strategy talk (ST) in any single session. When she was provided with data, she did not use the data, but instead, she justified what she was doing by describing her usual strategy on 63 occasions.

By contrast, S8 used data to the greatest extent. Her revisions were linked to feedback 42% of the time, although she said during the debriefing "But I think I'm not,

um, a terrifically model-based practitioner." She had the least irrelevant talk of all subjects; dialog and unrelated talk totalled 13 segments over the two sessions, and also the fewest strategy statements; a total of four over the two sessions. This preliminary comparison introduces a number of other questions. What other similarities and differences occur among the revisers? What are the differences between the revisions produced by these two subjects? Do their different styles give any clues as to the quality of their revisions? Examination of the segments not related to revision would provide some answers to these questions.

The question of most pragmatic interest to anyone involved in formative evaluation and revision is: how effective are the revisions? This has been partly answered in another study (Bordonaro, 1993), but further research will be needed to test the revisions produced or proposed by each of the different subjects. While this study did not attempt to evaluate the effectiveness of materials revised from input from a variety of sources, nor even whether these revisions were, indeed, improvements, it did examine the influence of personal skill, and the systematic use of heuristics and procedures by revisers.

This research has addressed an area which has not been the focus of extensive study, and the results should contribute to both the theory and practice of formative evaluation. The next stage would be to develop a model of reviser practice. This would form the basis of a heuristic to train novice revisers in this crucial stage of the development of instructional materials.

- Alexander, P. A., Shallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. <u>Review of Educational</u> <u>Research, 61</u> (3), 315-343.
- Anderson, J. R. (1993). Problem solving and learning. <u>American Psychologist</u>, <u>48</u>(1), 35-44.
- Andrews, D. H., & Goodson, L. A. (1980). A comparative analysis of models of instructional design. <u>Instructional Technology</u>, <u>3</u>(4), 2-16.
- Backler, M. (1990, March). Expertise analysis: Using cognitive tools to uncover how experts solve problems. Paper presented at NSPI Conference, Toronto.
- Baker, E. L. (1970). Generalizability of rules for empirical revision. <u>Audiovisual</u> <u>Communication Review</u>, 18 (3), 300 - 305.
- Baker, E. L. (1974). The role of the evaluator in instructional development. In G. D.
 Borich (Ed.), <u>Evaluating educational programs and products</u> (pp. 56-73).
 Englewood Cliffs, N.J.: Educational Technology Publications.
- Baker, E. L., & Alkin, M. C. (1973). Formative evaluation of instructional development. <u>AV Communication Review</u>, 21(4), 389-418.
- Beery, M. et al., (1981). Competencies for the instructional/training development professional. Journal of Instructional Development, 5 (1), 14-15.
- Bordonaro, T. (1993). <u>A comparison of the effectiveness. cost and efficiency of four</u> <u>formative evaluation conditions</u>. Unpublished master's thesis, McGill University, Montreal, Canada.
- Breuleux, A. (1991). The analysis of writers' think-aloud protocols: Developing a principled coding scheme for ill-structured tasks. In G. Denhière and J-P Rossi (Eds.), <u>Text and text processing</u>. North Holland: Elsevier Science Publishers, B.V.

Briggs L. J., & Wager W. W. (1981). <u>Handbook of procedures for the design of instruction</u>. Englewood Cliffs, NJ: Educational Technology Publications.

- Burkholder, B. L. (1981). The effectiveness of using the instructional strategy diagnostic profile to prescribe improvements in self-instructional materials. Journal of Instructional Development, 5(2), 2-9.
- Burt, C. W. (1989). Identification of critical variables in developmental testing and an experimental examination of the number and roles of participants in testing sessions. Unpublished doctoral dissertation, McGill University, Montreal.
- Cambre, M. (1981). Historical overview of of formative evaluation of instructional media products. Educational Communications and Technology Journal., 29(1), 1-25.
- Carey, L., Flower, L., Hayes, J. R., Schriver, K. A., & Haas, C. (1989). Differences in writers' initial task representations. (Technical Report No. 35). Pittsburgh, PA: Carnegie-Mellon University, Communications Design Center.
- Chi, M. T. H., Glaser, R., & Farr, M. J. (Eds.). (1988). <u>The nature of expertise</u>. Hillsdale, NJ: Erlbaum.
- Ciesla, J. L. (1976, April). <u>The effects of various formative evaluation procedures on</u> <u>instructional material revision in a large scale individualized science curriculum</u> <u>development project</u>. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.
- Clark, R. E. (1992). How the cognitive sciences are shaping the profession. In H. D. Stolovitch and E. J. Keeps (Eds.), <u>Handbook of human performance technology</u>. (pp 688-700). San Francisco: Jossey-Bass.
- Criswell, E. L. (1989). <u>The design of computer-based instruction</u>. New York: Macmillan.
- Davidove, E. A., & Reiser, R. A. (1991). Comparative acceptibility and effectiveness of teacher-revised and designer-revised instruction. <u>Educational Technology Research</u> and <u>Development</u>, 39(2), 29-38.

Debert, R. (1979). Revisionary tactics. Performance and Instruction, 18(7), 18-21.

- Dick, W. (1968). A methodology for the formative evaluation of instructional materials. Journal of Educational Measurement, 5(2), 99-102.
- Dick, W. (1980). Formative evaluation in instructional development. Journal of Instructional Design, 3 (3), 3-6.
- Dick, W., & Carey, L. (1985). <u>The systematic design of instruction</u>. (2nd ed.). Glenview,II.: Scott, Foresman.
- Dick, W., & Carey, L. (1990). <u>The systematic design of instruction</u>. (3rd ed.). Glenview, IL: Scott, Foresman.
- Dick, W., & Carey, L. (1991). Formative evaluation. In L.J.Briggs, K.L. Gustafson, &
 M.H. Tillman (Eds.) Instructional Design: Principles and applications (2nd ed.)
 (pp 227-267). Englewood Cliffs, N.J.: Educational Technology Publications Inc.
- Dick, W., & Reiser, R.A. (1989). <u>Planning effective instruction</u>. Englewood Cliffs, NJ: Prentice Hall.
- Dijkstra, S. (1991, April) <u>The instructional design for teaching to solve well- and ill-</u> <u>structured problems</u>. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.
- Duchastel, P. C. (1990). Cognitive designs for instructional design. <u>Instructional</u> <u>Science, 19</u> (6), 437-444.
- Dupont, D. (1980). <u>Développement et évaluation d'un modèle de révision de documents</u> <u>éducatifs</u>. Unpublished doctoral dissertation, Université de Montréal, Montreal.
- Dupont, D., & Stolovitch, H. D. (1983). The effects of a systematic revision model on revisers in terms of student outcomes. <u>NSPI Journal</u>, March 1983, 33-37.
- Duy, N. (1990). <u>Revising instructional materials: Uniformity among four revisers and</u> their attention to learner data. Unpublished master's thesis, McGill University, Montreal, Quebec.

- Ely, D. P. (1992). Trends in educational technology. In D. P. Ely,, & B. B. Minor, (Eds.), <u>Educational media and technology yearbook</u>, Vol 18, (1-30).
- Engler, D. (1976). Learner verification: A fine-grain analysis of go-go empiricism. Audiovisual Communication Review, 24(1), (5-20).

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

- Ericsson, K. A., & Simon, H. A. (1984). <u>Protocol analysis: Verbal reports as data</u>. Cambridge, MA: MIT Press.
- Ericsson, K. A., & Simon, H. A. (1993). <u>Protocol analysis: Verbal reports as data</u>. (Revised edition) Cambridge, MA: MIT Press.
- Faigley, L., & Witte, S. (1981). Analyzing revision. <u>College Composition and</u> <u>Communication, 32</u>, 400-414.
- Fenster, A. E., Harpp, D. N., & Schwarcz, J. A. (1990a). Artificial sweeteners. In A. E. Fenster, D. N. Harpp, & J. A. Schwarcz (Eds.) <u>The world of chemistry, Part 1</u> (pp. 67-71). (Available from McGill University Bookstore, Montreal, Quebec).
- Fenster, A. E., Harpp, D. N., & Schwarcz, J. A. (1990b). The diet-cancer relationship. In A. E. Fenster, D. N. Harpp, & J. A. Schwarcz (Eds.) <u>The world of chemistry</u>, <u>Part 1</u> (pp. 50-55). (Available from McGill University Bookstore, Montreal, Quebec).
- Fitzgerald, J. (1987). Research on revision in writing. <u>Review of Educational Research.</u> <u>57</u> (4), 481-506.
- Flaherty, E. G. (1984). The thinking aloud technique and problem solving ability. Journal of Educational Research, 68(6), 223-225.
- Flower, L., Carey, L., & Hayes, J. R., (1985). <u>Diagnosis in revision: The experts'</u> <u>opinion</u>. (Technical Report No. 27). Pittsburgh, PA: Carnegie-Mellon University, Communications Design Center.

- Flower, L., & Hayes, J. R. (1980). The dynamics of composing: Making plans and juggling constraints. In L. W. Gregg, & E. R. Steinberg (Eds.), <u>Cognitive</u> <u>processes in writing</u> (pp 31-50). Hillsdale, NJ: Erlbaum.
- Flower, L., Hayes, J. R., Carey, L., Schriver, K., & Stratman, J. (1986). Detection, diagnosis, and the strategies of revision. <u>College Composition and</u> <u>Communication, 37(1), 16-55.</u>
- Foshay, W. R., & Moller, L. (1992). Advancing the field through research. In H. D. Stolovitch and E. J. Keeps (Eds.), <u>Handbook of human performance technology</u>. (pp 701-714). San Francisco: Jossey-Bass.
- Gagné, R. M., & Briggs, L. J. (1979). <u>Principles of Instructional Design</u>. New York: Holt, Rinehart, & Winston.
- Gagné, R. M., Briggs, L. J., & Wager, W. W. (1988). <u>Principles of Instructional</u> <u>Design</u>. (3rd ed.). New York: Holt, Rinehart and Winston.
- Geis, G. L. (1986). Human performance technology: An overview. In <u>Introduction to</u> <u>Performance Technology</u>, Vol 1 (pp. 1-20). Washington, DC: NSPI.
- Geis, G. L. (1987). Formative evaluation: Developmental testing and expert review. <u>Performance and Instruction</u>, <u>26</u>(4), 1-8.
- Geis, G. L., Burt, C., & Weston, C.B. (1984, April). <u>Instructional development:</u> <u>Developmental testing</u>. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Genest, M., & Turk, D. C. (1981). Think-aloud approaches to cognitive assessment. In T. V. Merluzzi, C. R. Glass, & M. Genest (Eds.), <u>Cognitive assessment</u>. New York: New York University Press.
- Gilbert, T. F. (1992). Foreword. In H. D. Stolovitch and E. J. Keeps (Eds.), <u>Handbook</u> of human performance technology. (pp xiii-xviii). San Francisco: Jossey-Bass.
- Golas, K. C. (1983). Formative evaluation effectiveness and cost: Alternative models for evaluating printed instructional materials. <u>Performance and Instruction</u>, 22, 17-19.

- Graves, M. F., & Slater, W. H. (1991). A response to "Instructional texts written by five expert teams." Journal of Educational Psychology, 83(1), 147-148.
- Gropper, G. L. (1975). <u>Diagnosis and revision in the development of instructional</u> <u>materials</u>. Englewood Cliffs, NJ: Educational Technology Publications.
- Gustafson, K. L. (1991). <u>Survey of instructional development models</u> (2nd ed.). Syracuse, NY: Informations Resources Publications.

Hayes, J. R. (1989). The complete problem solver (2nd ed.). Hillsdale, NJ: Erlbaum.

- Hayes, J. R., Flower, L., Schriver, K. A., Stratman, J. F., & Carey, L., (1987)
 Cognitive processes in revision. In S. Rosenberg (Ed). <u>Advances in applied</u>
 <u>psycholinguistics</u> (Vol 2). Cambridge: Cambridge University Press.
- Hayes, J. R., & Simon, H. A. (1974). Understanding written problem instructions. In L.W. Gregg, (Ed.), <u>Knowledge and cognition</u>, Potomac, MD: Erlbaum.
- Henderson, E. S., & Nathenson, M. B. (1976). Developmental testing: An empirical approach to course improvement. <u>Programmed Learning and Educational</u> <u>Technology</u>, <u>13</u>(4), 31-42.
- Henderson, E. S., & Nathenson, M. B. (1977a). Case study in the implementation of innovation: A new model for developmental testing. In P Hills & J. Gilbert (Eds.), <u>Aspects of educational technology, X1</u> (pp. 114-120). London: Kogan Page.
- Henderson, E. S., & Nathenson, M. B. (1977b). Developmental testing: Collecting feedback and transforming it into revisions. <u>NSPI Journal</u>, <u>16</u>(3), 6-10.
- Israeloff, A. (1992). <u>Comparison of feedback generated by experts and learners during</u> <u>formative evaluation</u>. Unpublished master's thesis, McGill University, Montreal, Quebec.
- Jones, S. (1987). The analysis of depth interviews. In R. Murphy, & H. Torrance (Eds.), Evaluating education: Issues and methods. London, Paul Chapman

(originally published in R. Walker (Ed.), <u>Applied Qualitative Research</u>, Gower, 1985).

- Kandaswamy, S. (1980). Sequential model for appraising instructional superiority of revised materials. <u>Educational Communication and Technology Journal</u>, 28(3), 186-193.
- Kandaswamy, S., Stolovitch, H. D., & Thiagarajan, S. (1976). Learner verification and revision: An experimental comparison of two methods. <u>Audio-Visual</u> <u>Communication Review</u>, 24(3), 316-328.
- Kember, D., & Murphy, D. (1990). Alternative new directions for instructional design. <u>Educational Technology</u>, 30(8). 42-47.4
- Kemp, J. E. (1985). The instructional design process. New York, Harper, & Row.
- Knirk, F. G., & Gustafson, K. L. (1986). <u>Instructional technology: A systematic</u> <u>approach to education</u>. New York: Holt, Rinehart and Winston.

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

- Lefrere, P. (1981). Evaluation processes and strategies: Evaluating the process of evaluation. In F. Percival, & H. Ellington (Eds.), <u>Aspects of educational</u> <u>technology XV: Distance learning and evaluation</u> (pp. 127-133). London: Kogan Page.
- Le Maistre, C. (1991, April). <u>Revision: The missing link</u>. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Leshin, C. B., Pollock, J., & Reigeluth, C. M. (1992). Instructional design strategies and tactics. Englewood Cliffs, NJ: Educational Technology Publications.
- Macdonald-Ross, M. (1989). Towards a graphic ecology. In H. Mandl, & J. R. Levin (Eds.), <u>Knowledge acquisition from text and pictures</u>. (pp. 145-154). North Holland: Elsevier Science Publishers.
- Markle, S. M. (1989). The ancient history of formative evaluation. <u>Performance and</u> <u>Instruction</u>. <u>28</u>, 27-29.

- McAlpine, L. (1987). The think-aloud protocol: A description of its use in the formative evaluation of learning materials. <u>Performance and Instruction</u>, <u>26</u>(8). (18-21).
- McAlpine, L. (1990). Formative evaluation: Exploring the revision process and use of student data. Unpublished manuscript, McGill University, Montreal, Quebec.

McAlpine, L., & Weston, C. (1994). The attributes of instructional materials. <u>Performance Improvement Quarterly</u>, 7 (1). 19-30.

- McCormick, R. (1976). Evaluation of Open University course materials. <u>Instructional</u> <u>Science, 5</u>, 189-217. Amsterdam: Elsevier Scientific Publishing Company.
- Medley-Mark, V., & Weston, C. B. (1988). A comparison of student feedback obtained from three methods of formative evaluation of instructional materials. <u>Instructional</u> <u>Science</u>, <u>17</u>, 3-27.
- Mitchell, P. D. (1989). The future of education technology is past. <u>Canadian Journal of</u> <u>Educational Communication</u>, <u>18</u>(1), 3-27.
- Montague, W. E., Ellis, J. A. & Wulfeck, W. H. II. (1983). Instructional quality inventory: A formative evaluation tool for instructional development. <u>Performance</u> and Instruction, 22 (5), 11-14.
- Nathenson, M. B., & Henderson, E. S. (1980). Using student feedback to improve learning materials. London: Croom Helm.
- Nevo, D. (1985, March). <u>Experts' Opinion: A powerful evaluation tool</u>. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Newell, A., (1980). Reasoning, problem-solving, and decision processes: The problem space as a fundamental category. In R.Nickerson (Ed.), <u>Attention and</u> <u>Performance, Vol V111</u>, pp. 693-719. Hillsdale, NJ: Erlbaum.
- Newell, A., & Simon, H. A. (1972). <u>Human problem solving</u>. Englewood Cliffs, NJ: Prentice-Hall.

- Olson, J. R., & Biolsi, K. J. (1991). Techniques for representing expert knowledge. In K. A. Ericsson and J. Smith (Eds.) <u>Toward a general theory of expertise:</u> <u>Prospects and limits</u>. New York: Cambridge University Press.
- Piette, M. I., & Smith, N. M. (1991). Evaluating a Hypercard library instruction program. <u>Research Strategies</u>, 9(2), 87-94.
- Polanyi, M. (1958). <u>Personal knowledge: Towards a post-critical philosophy</u>. Chicago,IL: University of Chicago Press.
- Polanyi, M. (1959). The study of man. Chicago, IL: University of Chicago Press.

Polanyi, M. (1969). Knowing and being. London: Routledge and Keegan Paul.

Polyà, G. (1947). How to solve it. Princeton, NJ: Princeton University Press.

- Popham W. J., & Baker, E. L. (1971). <u>Systematic instruction</u>. Englewood Cliffs, NJ: Prentice-Hall.
- Rahilly, T.J. (1991). <u>An analysis of three learner-based formative evaluation conditions</u>.
 Unpublished master's thesis, McGill University, Montreal, Quebec.
- Rahilly, T. J., Weston, C.B., & McAlpine, L. (1991). <u>Issues in segmenting</u>. Unpublished manuscript, McGill University, Montreal, Quebec.
- Reitman, W. R. (1965). <u>Cognition and thought: An information-processing approach</u>. New York: Wiley.
- Reynolds, L., & Ehrlich, D. (1992, June). <u>Multimedia in industry and education: A</u> <u>decision model for design</u>. Paper presented at the European conference on educational research. Enschede, The Netherlands. (ERIC Document Reproduction Service No. ED 351 003).
- Richey, R. (1986). <u>The theoretical and conceptual basis of instructional design</u>. London: Kogan Page.
- Romiszowski, A. J. (1981). <u>Designing instructional systems:</u> <u>Decision making in course</u> <u>planning and curriculum design</u>. London: Kogan Page.

- Romiszowski, A. J. (1986). <u>Developing auto-instructional materials</u>. London: Kogan Page.
- Rosen, M. J. (1968). <u>An experimental design for comparing the effects of instructional</u> <u>media programming procedures: subjective vs. objective revision procedures</u>. Palo Alto, CA: American Institutes for Research in the Behavioral Sciences.
- Rothwell, W. J., & Kazanas, H. C. (1992). <u>Mastering the instructional design process</u>. San Francisco: Jossey-Bass.
- Rowland, G. (1992). What do instructional designers actually do? An initial investigation of expert practice. <u>Performance Improvement Quarterly</u>, 5(2), 65-86.
- Rowland, G. (1993). Designing and instructional design. <u>Educational Technology</u> <u>Research and Development, 41(1), 79-91.</u>
- Saroyan, A. (1989). <u>The review process in formative evaluation of instructional text: The</u> <u>role of content experts and instructional designers</u>. Unpublished doctoral thesis. Montreal, Canada. McGill University.
- Saroyan, A., & Geis, G. (1988). An analysis of guidelines for expert reviewers. Instructional Science, 17, 101-128.
- Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler (Ed.), <u>Perspectives</u> of curriculum and evaluation. (AERA Monograph Series on Curriculum Evaluation, No 1), (39-83), Chicago, IL: Rand McNally..
- Seels, B., & Glasgow, Z. (1989). <u>Exercises in Instructional Design</u>. Columbus, OH: Merrill.
- Shanteau, J. (1992) The psychology of experts: An alternative view. In G. Wright andF. Bolger (Eds.), Expertise and decision support. New York: Plenum.
- Simon, H. A., (1973). The structure of ill-structured problems. <u>Artificial Intelligence, 4</u>, 181 201.
- Simon, H. A., & Hayes, J. R. (1976). Understanding complex task instructions. In D. Klahr (Ed.), <u>Cognition and instruction</u> (pp. 269-285). Hillsdale, NJ: Erlbaum.

Simon, H. A., & Lea, G. (1974). Problem solving and rule induction: A unified view. In L. W. Gregg (Ed.) <u>Knowledge and cognition</u>. Potomac, MD: Erlbaum.
Smagorinsky, P. (1989). The reliability and validity of protocol analysis. <u>Written</u>

<u>Communication, 6</u> (4), 463-479.

Smith, P. L., & Ragan, T. J. (1993). Instructional design. New York: Macmillan.

- Smith, P. L., & Wedman, J. F., (1988). Read-think-aloud protocols: A new data source for formative evaluation. <u>Performance Improvement Quarterly</u>, <u>1</u> (2), 13-22.
- Steinberg, E. R. (1986). Protocols, retrospective reports, and the stream of consciousness. <u>College English</u>, 48(7), 697-712.
- Stolovitch, H. D. (1982). Applications of the intermediate technology of learner verification and revision (LVR) for adapting international instructional resources to meet local needs. <u>Performance and Instruction</u>, 21(7), 16-22.
- Stolovitch, H. D. (1983, April). Formative Evaluation Research: Overview and update. Presentation to the twenty-second 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: An introduction. In R. Beach & L.S.Bridwell (Eds.). <u>New</u> <u>directions in composition research</u> (pp 53-71). New York: The Guilford Press.
- Tessmer, M. (1993). <u>Planning and conducting formative evaluations</u>. London: Kogan Page.
- Tessmer, M. (1994). Formative evaluation alternatives. <u>Performance Improvement</u> <u>Quarterly</u>, <u>71</u>(1), 3-18.
- Thiagarajan, S. (1978). Instructional product verification and revision: 20 questions and 200 speculations. <u>Educational Communication and Technology Journal</u>, <u>26</u>(2), 133-142.
- Tiemann P. W., & Markle, S. M. (1984). On getting expertise into an expert system. <u>Performance and Instruction</u>, 23(9), 25-29.

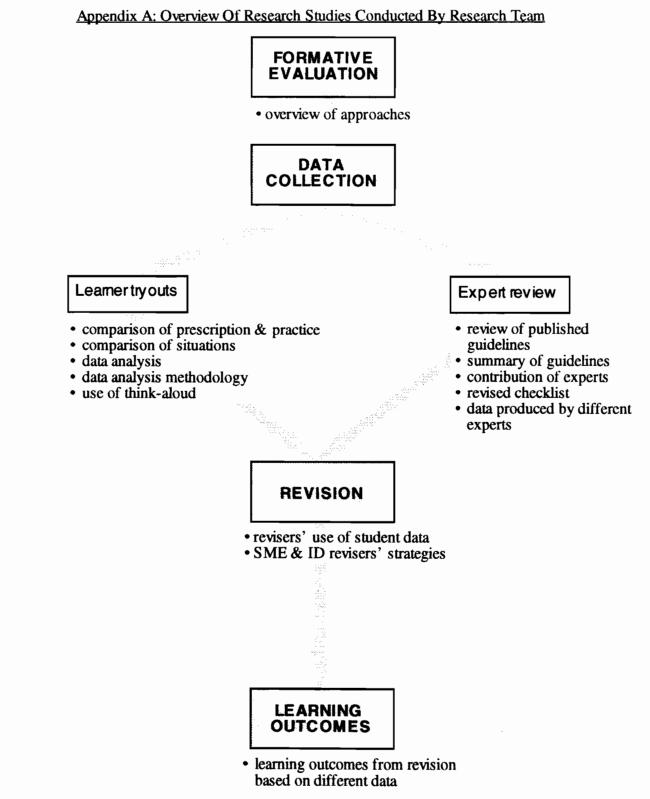
Tremblay, D. (in progress). <u>An analysis of the role of expert reviewers in formative</u> <u>evaluation</u>. Master's thesis, McGill University, Montreal, Canada.

- VanLehn, K. (1989). Problem solving and cognitive skill acquisition. In M. I. Posner (Ed.), Foundations of cognitive science (pp 527 - 579). Cambridge, MA: Bradford.
- Voss, J. F., & Post, T. A. (1988). On the solving of ill-structured problems. In M. T. H. Chi, R. Glaser & M. J. Farr (Eds.), <u>The nature of expertise</u> (pp. 261-285).
 Hillsdale, N.J.: Erlbaum.
- Wager, W. W. (1993). Instructional systems fundamentals: Pressure to change. <u>Educational Technology</u>, 33(2), 8-12.
- Waldron, J. S. (1973). <u>Instructional development unit: A guide to organizing an</u> <u>instructional development unit in health science</u>. Bethesda, MD: National Institutes of Health (DHEW). (ERIC Document Reproduction Service No. ED 125 625).
- Weston, C. B. (1987). The importance of involving experts and learners in formative evaluation. <u>Canadian Journal of Educational Communication</u>, <u>16</u>(1), 45-58.
- Willis, B. (1993). <u>Instructional development for distance education</u>: <u>ERIC Digest</u>. ERIC Document Reproduction Service No. ED 351007).
- Winn, W. (1990). Some implications of cognitive theory for instructional design. <u>Instructional Science</u>, 19, 53-69.
- Yelon, S. (1974). <u>Constructive evaluation: Improving large scale instructional projects</u>. New York, NY: Children's Television Workshop. (ERIC Document Reproduction Service No. ED 125 525).

APPENDICES

 \square

,



APPENDIX B

 \bigcirc

0

Contents of Packages Used by Administrators and Subjects

Script for Ice-breaker [1 (T1)]

Thank you for your participation in this study. I'd like to start by explaining what we're going to do today, and if you have any questions, feel free to interrupt me.

A bit of background on our research: This study is part of an on-going team research project concerned with the improvement of instructional materials through formative evaluation. My own piece of the research involves looking at revision and how experienced instructional designers carry out revision under different conditions. We'd like you to revise two modules from an introductory chemistry course for non-chemistry university students from both Arts and Sciences -- one today and one in the next session. The procedure for this session is outlined for you on the next sheet in this folder.

Script for Ice-breaker [1 (T2) w]

Thank you for coming back for this second session.

Let me remind you about our research: This study is part of an on-going research project concerned with the improvement of instructional materials through formative evaluation. We'd like you to revise another module from an introductory chemistry course for nonchemistry university students from both Arts and Sciences.

The procedure for this session is outlined for you on the next sheet in this folder, and this time it's slightly different.

Last time, you were asked to revise the module with no feedback on where problems had arisen during their testing. This time, I'll be giving you feedback data of various sorts that you can use to help you revise.

Script for Ice-breaker [1 (T2) n]

Thank you for coming back for this second session.

Let me remind you about our research: This study is part of an on-going research project concerned with the improvement of instructional materials through formative evaluation. We'd like you to revise another module from an introductory chemistry course for non-chemistry university students from both Arts and Sciences.

The procedure for this session is outlined for you on the next sheet in this folder, and this time it's slightly different.

Last time, you were asked to revise the module using the feedback we had obtained from experts and learners. This time, you're asked to revise another module, with no feedback provided to you.

Summary of Activities [2 (T1)]

1. I'll ask you to read a page of instructions, then check that you understand the task.

2. The technique you'll be using is called a Think-aloud. I'll give you some information on that next, and give you time to read it.

3. I'll give you a short warm-up task to allow you to practice thinking aloud.

4. I'll give you the module, ask you to work through it and revise it, and tape record your comments as you think aloud.

5. Finally, I'll ask you for some brief demographic information, and we'll discuss the materials more generally. I'll also answer any questions you may have after performing the task.

Is that short description of the procedure clear?

Before beginning, I have a consent form which I'd like you to read and sign, if you agree with what it says.

(Give consent form)

If you agree to the statements, would you please sign at the bottom of the form?

Summary of Activities [2 (T2) w]

1. I'll ask you to read a page of instructions, then check that you understand the task.

2. We'll again be using the Think-aloud process. I'll give you some information on that next, and give you time to read it, to refresh your memory.

3. I'll give you a short warm-up task to allow you to practice thinking aloud, if you feel you need it.

4. I'll give you the second module and the feedback we have from experts and learners, ask you to work through the module and revise it, and tape record your comments as you think aloud.

5. Finally, we'll discuss the materials more generally, and I'll answer any questions you have about the project.

Summary of Activities 2 [(T2) n]

1. I'll ask you to read a page of instructions, then check that you understand the task.

2. We'll again be using the Think-aloud process. I'll give you some information on that next, and give you time to read it to refresh your memory.

3. I'll give you a short warm-up task to allow you to practice thinking aloud, if you feel you need it.

4. I'll give you the second module, ask you to work through it and revise it, and tape record your comments as you think aloud.

5. Finally, we'll discuss the materials more generally and I'll answer any questions you may have about the project.

(Letterhead)

Consent form [T1]

I agree to participate in a programme of research conducted in the Department of Educational Psychology and Counselling at McGill University.

I give my consent to have the sessions audiotaped.

I understand that the data I provide will be coded to maintain confidentiality.

I understand that the data from this study may be published.

I understand that I can have a full description of the results of the study after its completion.

I understand that I am free to withdraw at any time during either session without negative consequences.

I understand that, at the end of the second session, I shall receive an honorarium of \$300 for my participation.

I HAVE READ AND UNDERSTOOD THIS CONSENT FORM AND I AGREE TO PARTICIPATE IN THE STUDY.

NAME:

SIGNATU	RE:
---------	-----

DATE:

This unit has been developed for use in a first-year chemistry course for both Arts and Science students who are not majoring in chemistry. The original draft was written by a subject-matter expert who is also a teacher. This person is dissatisfied with the effect of the materials on students. You are being brought in as a consultant to revise the unit. Please make **any** revisions that seem necessary to you. Feedback data has been collected from experts and learners, and I'll give you that data along with the original material.

Green-Pretest and Posttest scores from learners;

a copy of the tests;

oral comments, both positive and negative, from learners.

Yellow - Transcribed oral comments from a subject-matter expert;

Transcribed oral comments from an expert in the target population;

a summary of the qualifications of these experts.

I'll give you the module and the feedback data we have gathered on it, ask you to work through it, and tape record your comments as you think aloud while doing the revision.

While you are revising, we want you to speak your thoughts out loud so that we can record what you are saying. This technique is called a Think-aloud, which means saying out loud all the thoughts, questions, comments and strategies that go through your mind while you are performing a task. Since the Think-aloud process is unfamiliar to most people, we shall be having a short practice session in a few moments.

At times you may forget to think out loud, so my role will be to prompt you to continue. I shall also be keeping track of your comments. During the Think-aloud procedure, I can't answer any questions that you may ask, but there'll be a chance to do so at the end of the session. Please take your time on the task. I shall suggest times for taking breaks, but whenever you feel you want a break, please do not hesitate to say so.

Is there anything unclear about these instructions?

This unit has been developed for use in a first-year chemistry course for both Arts and Science students who are not majoring in chemistry. The original draft was written by a subject-matter expert who is also a teacher. This person is dissatisfied with the effect of the materials on students. You are being brought in as a consultant to revise the unit. Please make **any** revisions that seem necessary to you.

I'll give you the module and ask you to work through it, and tape record your comments as you think aloud while doing the revision.

While you are revising, we want you to speak your thoughts out loud so that we can record what you are saying. This technique is called a Think-aloud, which means saying out loud all the thoughts, questions, comments and strategies that go through your mind while you are performing a task. Since the Think-aloud process is unfamiliar to most people, we shall be having a short practice session in a few moments.

At times you may forget to think out loud, so my role will be to prompt you to continue. While you are speaking, I shall also be keeping track of your comments. In the Think-aloud procedure, I can't answer any questions that you may ask, but there'll be a chance to do so at the end of the session. Please take your time on the task. I shall suggest times for taking breaks, but whenever you feel you want a break, please do not hesitate to say so.

Is there anything unclear about these instructions?

The description of your task is very similar to what you read last time you were here, except that the conditions for the task are different this time, so would you please read this task description again?

This unit has been developed for use in a first-year chemistry course for both Arts and Science students who are not majoring in chemistry. The original draft was written by a subject-matter expert who is also a teacher. This person is dissatisfied with the effect of the materials on students. You are being brought in as a consultant to revise the unit. Please make **any** revisions that seem necessary to you. Feedback data has been collected from experts and learners, and I'll give you that data along with the original material.

Green - Pretest and Posttest scores from learners;

a copy of the tests;

oral comments, both positive and negative, from learners.

Yellow - Transcribed oral comments from a subject-matter expert;

Transcribed oral comments from an expert in the target population;

a summary of the qualifications of these experts.

I'll give you the module and the feedback data we have gathered on it, ask you to work through it, and tape record your comments as you think aloud while doing the revision.

While you are revising, we want you to speak your thoughts out loud so that we can record what you are saying. This technique is called a Think-aloud, which means saying out loud all the thoughts, questions, comments and strategies that go through your mind while you are performing a task. If you feel you need a warm-up exercise to remind you about the Think-aloud technique, I have one ready for you.

At times you may forget to think out loud, so my role will be to prompt you to continue. While you are speaking, I shall also be keeping track of your comments. In the Think-aloud procedure, I can't answer any questions that you may ask, but there'll be a chance to do so at the end of the session. Please take your time on the task. I shall suggest times for taking breaks, but whenever you feel you want a break, please do not hesitate to say so.

Is there anything unclear about these instructions ?

The description of your task is very similar to what you read last time you were here, except that the conditions for the task are different this time, so would you please read this task description again?

This unit has been developed for use in a first-year chemistry course for both Arts and Science students who are not majoring in chemistry. The original draft was written by a subject-matter expert who is also a teacher. This person is dissatisfied with the effect of the materials on students. You are being brought in as a consultant to revise the unit. Please make **any** revisions that seem necessary to you.

I'll give you the module and ask you to work through it, and tape record your comments as you think aloud while doing the revision.

While you are revising, we want you to speak your thoughts out loud so that we can record what you are saying. This technique is called a Think-aloud, which means saying out loud all the thoughts, questions, comments and strategies that go through your mind while you are performing a task. If you feel you need a warm-up exercise to remind you about the Think-aloud technique, I have one ready for you.

At times you may forget to think out loud, so my role will be to prompt you to continue. While you are speaking, I shall also be keeping track of your comments. In the Think-aloud procedure, I can't answer any questions that you may ask, but there'll be a chance to do so at the end of the session. Please take your time on the task. I shall suggest times for taking breaks, but whenever you feel you want a break, please do not hesitate to say so.

Is there anything unclear about these instructions?

NOW THAT YOU'VE READ THE INSTRUCTIONS, WOULD YOU PLEASE TELL ME YOUR UNDERSTANDING OF THE TASK?

WHAT IS THE "THINK-ALOUD" PROCESS ? [6]

Many people talk to themselves while performing a task. When you do this out loud, you are really performing a "think-aloud."

By recording the words you say as you think out loud, the researcher is able to study otherwise invisible mental processes that occur from moment to moment.

- As you revise the materials, say out loud whatever you are thinking.
- Don't worry about speaking correctly, stopping in the middle of thoughts or sentences, etc., but try to talk as continuously as you can while you are working.
- Try to avoid pauses. If we feel that you are not talking enough, we will prompt you.
- Don't try to explain or analyse what you are doing. Say only what is on your mind at the moment.
- Concentrate on the task you are given, and say aloud whatever occurs to you.
- Talk while you are reading, even if you are only skimming.
- As you write on the text, talk while you are writing.

While a think-aloud can't capture every thought that passes through your mind, there is a lot of evidence that is is an excellent method of viewing the processes that people go through while they are working. There is also evidence that thinking aloud does not interfere with your thinking during problem-solving, even though it seems very artificial at first.

In summary: relax, focus on the task you are given, and on whatever you are conscious of as you work. Say aloud everything that comes to mind.

Do you have any questions?

Warm-up instructions [7 (T1)]

Here is an activity to allow you to practice thinking out loud while you revise.

This is a section of a module similar to the one we'll ask you to revise in a few minutes. Please put yourself in the role of the consultant who has been asked to revise it, and make any revisions that seem necessary to you.

Think out loud while you are doing it.

In the actual revision task, you may take as long as you like, but for this practice, I'll stop you after about 3 minutes.

Warm-up instructions [7 (T2)]

Here is another activity to allow you to practice thinking out loud while you revise.

This is a section of a module similar to the one we'll ask you to revise in a few minutes. Please put yourself in the role of the consultant who has been asked to revise it, and make any revisions that seem necessary to you.

Think out loud while you are doing it.

In the actual revision task, you may take as long as you like, but for this practice, I'll stop you after about 3 minutes.

MONOSODIUM GLUTAMATE (MSG)

A 1969 rodent feeding study concluded that massive doses of MSG destroyed brain cells. This finding led to the voluntary elimination of MSG from baby foods by manufacturers which had been used to please the mothers' palates in the first place. The relevance of this study to humans, however, is very questionable since a number of subsequent studies with primates showed no effect upon jnjection or force feeding of MSG.

The general conclusion based upon the scientific evidence is that some people show an idiosyncratic response to MSG when they are exposed to high doses on an empty stomach. Symptoms may vary, and are not reflected by objective measurements or by blood levels of glutamate. In rare cases, in suffite sensitive individuals, large amounts (5-10 grams) of MSG may trigger asthmatics attacks. As far as the general population is concerned, the percentage of people subject to Chinese Restaurant Syndrome, perhaps 2%, is lower than that of people who have allergic reactions to other foods. Although MSG obviously contains sodium, it only contributes about 1% of the total intake of this mineral in the average diet.

Today, glutamate research is focused on the role of the substance as a neurotransmitter. It is believed that about half the nerve cells in the brain use glutamate as a chemical communicator and that the brain damage caused by a lack of oxygen, as in a stroke, is actually due to the release of excessive amounts of glutamate triggered by oxygen starvation. This release causes overstimulation of brain cells, leading to their death. The search is on for drugs which can block glutamate receptor sites on nerve cells and thus help reduce the damage. It is of interest that compounds closely related to the street drug phencyclidine (PCP) have actually shown some potential in this regard. There is even hope that certain diseases of the nervous system, such as Parkinson's and Alzheimer's may respond to glutamate blocking therapy.

Although excessive levels of glutamate in the brain may prove damaging, nobody is suggesting that these levels can be caused by dietary intake. As we have seen, even the allegation that the "Chinese Restaurant Syndrome" is caused solely by monosodium glutamate is on rather shaky scientific ground. Indeed, if this were the case, and given the fact that mushrooms and tomatoes are high in naturally occurring glutamate, we should be witnessing an epidemic of Chinese Restaurant Syndrome among pizza eaters!

[8(T1)]

PRESERVATIVES

There is no question that contamination by microorganisms is the major cause of food related illness in the world. Millions of cases of diarrheal disease are caused by the rapid growth of Salmonella bacteria in the intestinal tract which are introduced by contaminated foods. There is extensive suffering from the ingestion of toxins produced by Staphylococcus bacteria in spoiled food and there are deaths from botulism poisoning.

Some effects of tainted food are more insidious. Some molds, such as those found on improperly stored peanuts, produce trace amounts of carcinogenic aflatoxins. Certain fruits and vegetables degrade when their tissues are damaged while oxygen in the air can also contibute to food deterioration by promoting rancidity in fatty foods. Fortunately, a number of physical and chemical processes are available to reduce the risk of food spoilage and contamination.

Antimicrobial agents

Certain chemicals can inhibit, although not kill, a variety of fungi, yeasts, molds and bacteria. Sodium benzoate, which occurs naturally in cranberries and prunes is effective in controlling the growth of yeast and bacteria and is used extensively in beverages, jams, jellies, pickles and margarine. Sodium or calcium propionate in turn is effective against molds and bacteria but not yeasts. It is therefore an ideal substance for the prevention of contamination in bread after it leaves the oven. Undoubtedly, however, its presence alarms some people who do not like the idea of "chemical additives" in their bread.

[8(T2)]

THE DIET CANCER RELATIONSHIP

[10 (M1)]

"Sound nutrition is not a panacea. Good food that provides appropriate proportions of nutrients should not be regarded as a poison, a medicine, or a talisman It should be eaten and enjoyed." This statement by the Food and Nutrition Board of the National Research Council in the U.S in a 1980 publication called "Toward Healthful Diets" raised more than a few eyebrows. Reaction from consumer groups was furiously negative. These groups along with many individuals objected to the conclusion that no specific dietary advice was appropriate for all citizens. The recommendation of a balanced diet with moderation in consumption did not sit well with people who were convinced that a great many of the ills of North American society are related to improper nutrition. A document detailing the evils of food additives, the benefits of vitamin supplementation and the virtues of "organic" foods would undoubtedly have received more favorable reaction. Science however cannot deal with emotions, beliefs or anecdotal evidence; it must be based on facts stemming from well controlled and reproducible experiments. Unfortunately in the area of nutrition it is very difficult to design and carry out studies which lead to conclusive results. Accordingly many reports of results are speckled with phrases like "may cause", "is consistent with", "is associated with"; all of which imply uncertainty. The difficulty of providing "proof" one way or another in the areas of food science and nutrition leaves the door open to a variety of opinions not only among the alarmists and self styled authorities but among nutritional experts as well.

Indeed, just two years after the above mentioned report the National Research Council issued a new document entitled "Diet, Nutrition and Cancer" with more specific recommendations reflecting the state of knowledge and information pertinent to the diet and the incidence of cancer. The guidelines now recommended a reduction of fat intake from about 40% to 30% of total calories, a reduction in the consumption of cured, pickled and smoked foods and an increase in the consumption of whole grain ccreal products as well as fruits and vegetables, especially those rich in carotene. Vegetables belonging to the cabbage family were highly recommended but vitamin supplementation was not advised. The new report was in turn also criticized. Many scientists believe that not enough is known about the diet-disease connection to warrant specific guidelines for the population as a whole and furthermore the suggestion was made that if the guidelines were improperly applied they could lead to nutritional deficiencies. In light of the ongoing controversy it is appropriate to examine the studies and the kind of data that have lead to the debated recommendations. An examination of this controversy also serves to underline the need for a basic scientific understanding of chemical and nutritional concepts. Familiarity with terms like "minerals", "vitamins", "fat", "fiber", "carotene" etc. is essential for an objective and critical discussion of the relationship between diet and cancer.

There appears to be little doubt that many cancers are environmentally related. Epidemiological studies have clearly shown large differences in cancer rates between countries. For example, breast and colon cancer rates in many areas of the world are less than one fifth that in North America. The Japanese in turn have the highest incidence of stomach cancer in the world. Immigrants from other countries to the U.S. and Canada however experience the local cancer rates, suggesting an environmental influence.

Perhaps the best demonstration of this environmental effect comes from a study made public in 1984 by the National Cancer Research Institute of Japan. An epidemiological study spanning 16 years and involving over 100,000 men clearly showed that the incidence of cancer was greatest among those who smoked, drank alcohol, ate meat regularly and did not consume vegetables daily. Indeed the absence of vegetables from the diet appeared to increase the risk of a wide variety of cancers. The results of the survey are summarized below:

RELATIVE MORTALITY RATES

Smoking	Drinking	Meat	Vegetables	
No	No	No	Ycs	1.0
No	Yes	Ycs	No	1.1
Yes	Yes	Yes	Yes	1.7
Yes	Yes	No	No	1.8
Yes	No	Yes	No	1.8
Yes	Yes	Yes	No	2.5

The protective effects of vegetable consumption are dramatically illustrated by the above data; in fact even in the high risk group (smokers, drinkers and meat eaters) the risk of cancer can be reduced by one third if vegetables are regularly eaten. This protective effect may be manifested through the fiber, Vitamin C or carotene components of the vegetables as discussed below.

Accordingly many cancer experts now estimate that as much as 90% of North American cancers are environmentally determined and that a large fraction of these should therefore be avoidable. "Environmental" must not be confused with "man made"; in the present context the word is used to differentiate from "genetic" factors. Cigarette smoking and toxic wastes are environmental and obviously "man made", but exposure to sunlight and the consumption of naturally occurring carcinogens can also be termed "environmental". In fact, Bruce Ames of the University of California (Berkeley) has concluded after a survey of the scientific literature that most of the carcinogens that non-smokers encounter in their daily life come from natural foods and cooking methods. For example celery and parsley contain a carcinogen which becomes activated by light; mushrooms, beans and even alfalfa sprouts contain compounds which may increase the risk of cancer. Cooking, especially when food is browned or burned adds carcinogens to the diet. On the other hand, suggests Ames, food also appears to contain natural anti- carcinogens like Vitamins C and E, selenium and carotene which may decrease the risk of the dreaded disease. The fact that cancer rates aside from those related to smoking have remained almost constant over the years appears to imply that the "natural" components of the environment may be more important than the "man made" factors in inducing cancer.

In a controversial article in Science, 221, 1256 (1983), Ames produced summarized the many natural foods (above) which contained various carcinogens. In this same article, he also indicated that there were many foods which were also anti-carcinogens. The main idea here was that a minimum of the questionable foods coupled with a reasonable amount of the "good" ones (vide infra) would provide as good a balance of risk/benefit as could be achieved in this very complex area. Ames was criticized by a group of 18 academics, union officials and environmentalists in a 1984 letter to Science for "trivializing" cancer risks. Ames recently published a summary of relative risk factors for cancer by a careful (but controversial) examination of the literature. The resulting index called HERP (Human Exposure dose/Rodent Potency dose). This index considers two questions: How much of the material causes considerable rates of cancer in lab animals, and how much of it might an average person be exposed to over a lifetime? The rankings do not predict a person's actual chances of developing cancer, but show comparisons. If the relative ranking of tap water is 1.0, then peanut butter (2 tablespoons/day) is 30

(aflatoxin risk) as is comfrey tea (1 cup/day) (symphytine, a natural pesticide is present). One pack 181 /day of cigarettes is rated at 12,000 while the risk of cancer from PCBs (once used in electrical transformers) is 0.2. Needless to say, such a detailed list has created concern and discussion and will stimulate research in the future.

Since the second World War some 50,000 synthetic chemicals have been introduced into the environment with about 500 new ones coming into use every year. Many of these are mutagenic or carcinogenic in lab tests yet the cancer epidemic that many scientists (even Bruce Ames at one time) have predicted has not materialized. Accordingly there is widespread, though certainly not universal, belief that most cancers are caused by natural carcinogens. Many of these carcinogens are produced by plants as natural pesticides to ward off insects. Ironically the current practice of breeding insect resistant plants in order to minimize the use of synthetic pesticides may actually be introducing new carcinogens into the diet. It is also a fact of course that not everyone gets cancer even though everyone consumes natural carcinogens. The explanation for this apparent inconsistency may lie in the possibility that whereas chemicals isolated from food can cause cancer, the whole food does not. Mutagens and "anti- carcinogens" are often present in the same food. For example the potentially harmful effects of the psoralens in parsley and celery may be counteracted by the carotene and vitamin components of these foods. It appears then that attention to a scientifically balanced diet may be more important in warding off cancer than worrying about the trace amounts of synthetic carcinogens in the environment. The following summarizes the current state of knowledge in this important area.

The Dietary Fat-Cancer Relationship

The above mentioned recommendation to reduce the fat content of the diet stems mostly from correlations noted by epidemiologists. A strong correlation exists between per capita fat intake and breast cancer mortality in women as well as between fat intake and mortality from colon cancer. It must be pointed out however that such associations do not imply cause. For example a similar correlation exists between gross national product and breast cancer. Although the "per capita" correlation of dietary fats with cancer is strong, there appears to be no conclusive correlation of individual fat consumption and cancer. There may be other variables in the relationship as well. Hormones like estrogen have been linked with cancer. Does the fact that women are having fewer and later pregnancies influence the average estrogen levels ? Could it be the added calories and not the fat per se which is instrumental? The human feeding studies which would be needed to clarify the situation can never be ethically done but studies in animals do suggest that higher levels of fat intake cause mammary tumors. Theoretically the argument can be put forward that fats cause cancer by undergoing oxidation in cells leading to the production of cancer causing reactive species called free radicals. These free radicals then damage the DNA of the cell, leading to improper replication. If this mechanism is correct, unsaturated fats may pose a greater risk since they are more easily oxidized. Some studies have indeed shown an association between cancer and "trans" fatty acids which are produced when vegetable oils are converted into margarine. Adequate Vitamin E, beta carotene and selenium consumption may prevent the oxidation of fats.

Dr. Keith Ingold at the National Research Council in Ottawa has in fact shown that Vitamin E is the major "free-radical trapping" anti-oxidant in human blood. Beta carotene can also act as an antioxidant, especially at low oxygen concentrations such as are found in cells. It is noteworthy that this important research started out as an investigation into why engine oils break down upon exposure to oxygen in the car's engine; a nice demonstration of how important results can come from seemingly "unimportant" research. Similarly the antioxidants BHT and BHA which had originally been developed to prevent fats in cereals from going rancid (and incidentally have been much maligned) may turn out to have an important role in not only the prevention of cancer but in actually slowing down the aging process.

Colon cancer has also been associated with high fat, high cholesterol diets. Once again though, epidemiological studies in individuals have yielded inconsistent results. Animal feeding studies in turn have shown that dietary fat promotes colon cancer. Furthermore, populations with high rates of colon

cancer have increased levels of bile acids in the feces; these have been associated with cancer and are 182 known to be formed in larger amounts in high fat, high cholesterol diets. In summary, the evidence may appear to be somewhat circumstantial, but the recommendation to reduce fat content by 25% does not represent a risk as long as a balanced diet is maintained.

The Cured Foods-Diet Association

Once again population studies have shown that cancers of the stomach and esophagus are more common in countries such as China, Japan and Iceland where the diet is high in foods that are salt cured and smoked. There is no doubt that smoke contains cancer causing compounds and salt has been reported to promote gastric cancer in rats. Sodium nitrite, a pickling agent and preservative used in cold cuts, hot dogs, ham, etc. has been linked with the potential formation of nitrosamines, known carcinogens, in the body. Based upon these observations, limiting the intake of such cured or smoked foods would appear to be wise. Yet, even this recommendation has been challenged. It has been pointed out that the death rate from stomach cancer has been declining in North America while the consumption of processed meats has been rising. Furthermore, nitrite addition is so strictly regulated now that only minimal amounts are used; in fact the amount of nitrite now added can only prevent growth of the clostridium Botulinum organism if it is used in conjunction with salt. It is also true that most of the "smoked" foods presently marketed are smoked with liquid smoke. This is made by passing smoke through water; since the carcinogenic compounds do not dissolve in water foods "smoked" by this process are safer than "naturally" smoked foods. Although credence can be given to these criticisms, it must also be pointed out that foods high in smoke flavor and nitrites are generally high in fat and thus in calories-perhaps enough of a reason to minimize consumption.

The Selenium-Cancer Association

Selenium is a mineral required by the body in "trace" amounts. It plays a role in the activity of the enzyme glutathione peroxidase, an enzyme which protects cells from damage by oxidation. Consistent with this activity is the observation that mammary cancer in rats fed a high polyunsaturated fat diet can be inhibited by selenium. Selenium is found in the soil and is absorbed by crops. High soil selenium areas correlate inversely with cancer but these areas are also less populated and differ from low soil selenium areas in several respects. Indeed lung cancer rates are lower in countries where tobacco contains more selenium. Mexican and Colombian tobaccos have three times as much selenium as American and British tobaccos. Some correlations between blood selenium levels and cancer have also been noted and preliminary research has shown that the selenium content of hair and nails may reflect blood levels. High intake of selenium can be toxic and the presently available information does not warrant the recommendation of supplements.

The Cancer- Vitamin C and E Connection

The evidence for this association is essentially anecdotal although both of these vitamins are antioxidants and therefore could behave as anti- carcinogens. Vitamin E has been reported to reduce mutations in some bacterial systems and Vitamin C does block the conversion of nitrites to nitrosamines. For the latter reason Vitamin C is added to hot dogs. Similarly since both tomatoes and lettuce contain Vitamin C they can conceivably do more than just dress up the appearance and flavor of a bacon sandwich. Indeed a BLT may be the best way to consume bacon. There is however no evidence that either Vitamin E or C can prevent cancer.

The Cancer-Vitamin A Connection

Remember the stories about eating carrots to see better? This may be stretching the point, but the vitamin A in carrots does play an essential role in the chemistry of vision. Furthermore, the vitamin and its precursor compound (beta- carotene) may also protect the body against cancer. The rationale for this belief lies in the fact that vitamin A plays an important role in the control of cell differentiation and in that both vitamin A and especially beta-carotene are efficient scavengers of chemical species called free radicals. Since loss of cell differentiation is a basic feature of cancerous cells and since free

radicals are unstable, highly reactive chemicals which can damage our genetic materials (DNA and RNA) there is good reason to suspect that these two nutrients may have a protective effect against cancer.

Vitamin A itself can be obtained from animal products such as liver, eggs and meat or it can by synthesized by the body from. beta carotene. Many green vegetables produce this bright orange compound but the richest sources are pumpkins, spinach and of course carrots.

In 1975 a major epidemiological study showed that Norwegian men consuming more than the average amount of vitamin A had less than half the rate of lung cancer as compared with men having below average consumption of the vitamin. Similar findings were also reported in the following 5 years from scientists in Japan, Singapore and the United States.

A further study (Nov. 1981) published in the British medical journal Lancet supported the hypothesis that the pro-vitamin A (beta carotene) and not the vitamin itself was the beneficial factor. The study showed that there was an inverse relationship between intake of dietary beta-carotene and lung cancer in 1,954 middle aged male smokers over a period of 19 years. Intake of preformed vitamin A did not show a significant effect.

Unfortunately, studies on vitamin A are often limited due to its toxicity. High levels of vitamin A lead to liver damage, headaches, lack of appetite, hair loss, menstrual problems and retarded growth in children -- problems sometimes seen among vitamin and health food faddists. On the other hand, optimal investigative approaches are possible with beta carotene since there are no known serious side effects, even with doses so high as to cause and obvious orange skin coloration. In recent years synthetic analogs of vitamin A have been prepared in an effort to reduce its toxicity. These safer compounds are now being tested with high risk groups to determine if other forms of cancer can be prevented. One such group consists of albino children in Africa who have a 100% risk of developing skin cancer. In addition, at the present time the U.S. National Institute of Health has invited all male physicians between the ages of 40 and 85 to participate as subjects in a placebo-controlled general study of beta- carotene and cancer.

A major report on this issue published in the New England Journal of Medicine, March 1984 (by the Harvard School of Public Health) explained that although the protective effect against lung cancer of beta-carotene is strongly supported by many studies, there are indications that these effects may not apply to other types of cancer.

In conclusion, it should be noted that the main cause of lung cancer, smoking, also increases one's risk of several other serious diseases, including atherosclerosis -- a primary cause of death in North America. However, there is no evidence that either vitamin A or beta carotene affects this condition in any way.

The Cancer-Fiber Connection

Roughage? Unappetizing, tasteless, completely indigestible but... it fights cancer!

It all started with Dr. Dennis Burkitt's 20-year observation of diets and incidence of colorectal cancer in rural Africa. The British surgeon noted that although cancer of the lowest five to six feet of the intestine is very prevalent in the western world it is almost nonexistent among people in Africa consuming a high fiber diet. In Canada, about 100,000 people get colon cancer every year, half of whom die within the same year. The same high frequency of this malignancy has been found in the U.S., Scotland, Denmark and especially New Zealand, countries which consume the highest amounts of meat and animal fat around the world.

The incidence of this type of cancer appears to be 100 times more prevalent in the lowest 1% of the small intestine. This leads scientist to believe that carcinogens are not swallowed with our food but are produced in the colon from material in the feces. It has been suggested that bile acids (biomolecules

naturally released into the gut in response to the presence of fat in the diet) are chemically altered by bacteria to produce carcinogens. High colon cancer areas have been found to be much more abundant in colorectal cancer patients than in control groups. In a recent study conducted by Dr. Tracy Wilkins, a microbiologist at the Virginia Polytechnic Institute in Blacksburg, a chemical mutagen, named faecapentaene, was isolated from the feces of about 20 per cent of the white residents of Johannesburg. The same compound was detected in less than 2 per cent of the rural population. The diet of the urban community is very similar to ours (high in refined carbohydrates and fat), whereas that of the rural population is low in meat and fat and high in fruits and vegetables. Although most carcinogens are mutagens not all mutagens are carcinogens, and therefore the presence of faecapentaene does not necessarily mean that it is the cause of cancer. Dr. David Kingston, a chemist at the Virginia Polytechnic Institute, has synthesized this compound and its cancer-causing potential will now be investigated in laboratory animals.

These findings certainly support the theory that fiber, which increases the rate of feces elimination, should lower one's chances of developing cancer of the colon. However, there are some inconsistencies in the findings related to the effects of fiber. For instance, in a Canadian study published in 1980 higher consumption of dietary fiber was shown not to have any significant effect on cancer whereas in Puerto Rico high consumption was associated with higher incidence of colon cancer. Such discrepancies may be related to the extremely heterogeneous nature of dietary fiber. Dietary fiber is a mixture of indigestible chemicals: cellulose, hemicellulose, lignin and pectin. Preliminary studies have shown that wheat bran and fiber from citrus fruits protect laboratory animals against chemically-induced colon cancer. Since citrus fruits are also an excellent source of vitamin C (a scavenger of carcinogenic free radicals) an orange a day, or even the traditional apple a day, may not be such a bad idea.

GUIDELINES FOR AN ANTICANCER MENU

-decrease consumption of fats, nitrite-cured meats, smoked or charcoal-broiled meats and large amounts of alcohol.

-increase consumption of foods rich in dietary fiber,. beta carotene, vitamins A, E and C and the mineral selenium (megadoses of dietary supplements are presently not recommended).

-consume often, cruciferous vegetables such as cabbage, broccoli, Brussels sprouts and cauliflower.

RECENT REVIEW

A recent summary which gives a balanced report is from Scientific American, November, 1987, p. 42.



ARTIFICIAL SWEETENERS

[10 (M2)]

The replacement of sugar by artificial sweeteners in the diet may be of importance in weight reduction, in the maintenance of dental health and in making available a greater variety of foods for diabetics. There are three such sweeteners allowed in Canada today. Aspartame (trade names Equal and NutraSweet) is allowed as an additive in a large number of foods and also as a tabletop sweetener, while cyclamate (Sweet & Low) and saccharin (Hermestas) are allowed only as tabletop sweeteners. In the U.S. cyclamate has been banned since 1970, but saccharin and aspartame can both be used as

Saccharin, the cyclamates and aspartame were all discovered through lucky accidents. In 1879, Constantine Fahlberg, a chemist working at Johns Hopkins University in Baltimore, linked an unusual sweetness in a slice of bread he was eating to a chemical residue on his hands from the lab. Eventually he traced the sweetness to "benzoyl o-sulfonamide" which was given the name "saccharin." Fahlberg soon became wealthy but his reputation suffered greatly when he applied for patents in secret and refused to give any credit to his research director, Ira Remsen, under whose guidance he was working when the serendipitous discovery was made. In 1937, Michael Sveda, working on antibacterial substances at the University of Illinois, noted a sweet taste on a cigarette he had placed on the bench top and discovered the sweetening power of cyclamate. Similarly, aspartame was discovered in 1965 when James Schlatter, a researcher at Searle Laboratories, licked his finger while engaged in work on anti-ulcer medications.

additives or as tabletop sweeteners. The U.S. Food and Drug Administration has recently given

approval to a new sweetener, acesulfame K, which could be in common use by 1989.

Artificial sweeteners are among the most controversial food additives due to allegations of adverse health effects. These allegations include dermatological problems, headaches, mood variations, behavior changes, respiratory difficulties, seizures and cancer. A very large number of studies on these substances have been carried out with conclusions ranging from "safe under all conditions" to "unsafe at any dose." Scientists are divided on the issue of artificial sweetener safety, and in many cases researchers seem to arrive at results which confirm their views. In scientific, as well as in lay publications, supporting studies are often widely referenced while opposing results are de-emphasized.

The general conclusion is that sweeteners when used in modest amounts represent a minimal risk. Abusive amounts, which are not rare, are associated with increased, but still low risk. Most importantly, however, the risk-benefit ratio of artificial sweeteners is unclear. A number of studies have shown that the use of sweeteners is not associated with weight loss. Apparently, people saving on calories by using artificial sweeteners are so proud of themselves that they compensate by indulging more extensively in other foods. The elimination of sugar from soft drinks is beneficial in terms of the dental health of children, but it would make more nutritional sense to greatly reduce the over 40 gallons per person per year rate of consumption of soft drinks altogether. Artificial sweeteners do afford real benefits to diabetics by allowing them to eat foods which otherwise would have to be severely restricted.

SACCHARIN

Saccharin is about 300 times sweeter than sugar and can be called a true "non-caloric" sweetener since it is eliminated by the body intact. It does have a disagreeable aftertaste which can be minimized by the addition of a small amount of the amino acid, glycine, or by combining it with cyclamate. Sugar contributes not only sweetness but also "body" to a beverage. When sugar is replaced by an artificial sweetener, thickeners such as gum arabic, or carboxymethyl cellulose are often used to impart a proper mouthfeel.

In 1972, saccharin was removed from the "Generally Recognized as Safe" list in the U.S. when experiments showed a possible link with bladder cancer in rats. This meant that the additive was now

68

subject to the "Delaney Clause," which had been part of an amendment to the Federal Food, Drug and Cosmetics Act in 1958. This clause states that any additive which in any dose causes cancer in any species of animal has to be immediately banned. However, the accusatory evidence was judged to be inconclusive and saccharin use was allowed until further research either confirmed or refuted the evidence. Follow up studies on rats, mice, hamsters and monkeys were negative except for one case in which there was an increased incidence of bladder cancer in male rats fed high lifetime doses of saccharin if, and only if, their parents had also been fed similar high lifetime doses. These doses made up 5% of the rats diet and were the equivalent of a human consuming 800 cans of diet drink per day. This of course is a very large amount, but the technique of feeding large doses to rats over a short time span is a scientifically accepted method of estimating what may happen in humans exposed to small doses of the substance over a long period.

When Canadian studies in 1977 showed that the bladder cancer was indeed caused by saccharin and not an impurity as some had suggested, the Canadian government banned the use of saccharin as an additive. In the U.S. public and industry pressure forced Congress to pass the "Saccharin Study and Labelling Act," which put a moratorium on a ban as would have been dictated by the Delaney Clause. This moratorium has been repeatedly extended. Since saccharin costs only about one twentieth as much as sugar, and since it allows for a greater variety of products and hence more shelf space, it is easy to see why the industry would oppose any restriction on the use of this high profit substance.

The cancer case against saccharin is by no means iron clad. A theoretical rational for cancer induction is difficult, since saccharin passes through the body unaltered. The suggestion has been made that saccharin exerts its effects by temporarily attaching to and disrupting certain proteins. As mentioned above, only "double generation" studies in male rats have confirmed bladder cancer, and even in these cases the incidence of tumors was seen to decline sharply with dose. The really worrisome aspects of carcinogenesis, that is tumors in several animal species, rapid tumor formation, and tumors in various organs, have never been associated with saccharin. Furthermore, human "case-control" studies have examined the lifestyle and dietary intakes of a large number of bladder cancer patients without revealing any link to saccharin. Neither have any studies found increased incidence of tumors among diabetics, a population subgroup which would be expected to be at increased risk due to extensive use of artificial sweeteners. It is likely then, that any but a vanishingly small effect of saccharin on human bladder cancer would have been detected. The possibility exists that saccharin is a "co-carcinogen," that is a substance which may initiate cancer in cells which have already been exposed to some carcinogen. In this context, it should be pointed out that there are many naturally occurring substances in our diet, for example the amino acid tryptophan, which are known to promote cancer at high doses but represent no risk at all as part of a normal diet. In all likelihood, this is also the case with saccharin. Calculations based on the available animal models show a "worst case scenario" of a one in a million increased risk of bladder cancer due to the consumption of two saccharin sweetened beverages a day over a lifetime.

CYCLAMATES

Three similar compounds, namely sodium cyclamate, calcium cyclamate and cyclamic acid are collectively referred to as the "cyclamates." They are about thirty times sweeter than sugar and are chemically more stable than saccharin or aspartame. Some of a cyclamate dose is excreted by the body unchanged, but some is converted to a cyclohexylamine, a compound whose safety has been questioned. Like saccharin, cyclamates were granted "Generally Recognized as Safe" status in the U.S. in 1958 based on years of apparent problem free use. This status was rescinded in 1969 when bladder tumors were discovered in rats which had been fed a cyclamate-saccharin (10:1) mixture. Since saccharin was a minor component of this mixture, cyclamates were implicated in the tumor formation.

The 1969 study by the independent Food and Drug Research Laboratories was widely criticized, and never reproduced with a statistically significant number of animals. In fact, more recent studies on rats, mice, dogs, hamsters and monkeys, using both cyclamates and their breakdown product,

cyclohexylamine, have shown no effect on cancer rates. Indeed, unlike saccharin, the question of whether cyclamates are carcinogenic or co-carcinogenic at all, has not been adequately answered. Neither has the allegation that cyclamates are mutagenic, that is, capable of producing genetic damage which can be inherited, been substantiated.

The only reproducible effect of cyclamates has been testicular atrophy in rats when fed large doses. This problem, which has not been seen in humans, is caused by the metabolic product cyclohexylamine. Taking into account the maximum dose at which no such effect is observed in rats, and building in a hundred fold safety factor, it is possible to arrive at an "Accepted Daily Intake" of cyclamate which represents an insignificant risk. Accordingly, the Health Protection Branch in Canada allows cyclamate as a tabletop sweetener. The U.S. ban on cyclamates is still in effect, but may be rescinded in light of the fact that the Food and Drug Administration has reviewed the large number of studies submitted to it in the 1980's and has concluded that neither cyclamates nor cyclohexylamine are carcinogens.

ASPARTAME

Aspartame cannot be considered a "non-caloric" sweetener since it is broken down in the digestive tract into its components which are absorbed and metabolized. These components, aspartic acid, phenylalanine and methanol account for the 4 calories per gram energy rating of aspartame. However, since the substance is about 180 times sweeter than sugar, very little needs to be used in foods and beverages to achieve a satisfactory degree of sweetness. Diet drinks normally contain about 60 milligrams per 100 milliliters, which translates to roughly 200 milligrams per serving. Industry figures show that 99% of aspartame users consume less than 34 mg per kg of body weight per day. For a seventy kg person this means less than 2.38 grams. The average consumption is about 500 milligrams per day. Aspartame cannot be used in cooked or baked foods since it breaks down into its components and loses its sweetening power.

Aspartame is the most widely researched food additive to have ever come on the market. As with any other newly introduced substance, reports of adverse reactions were expected since no amount of testing can preclude idiosyncratic reactions in a small minority of the population. In reality, the number of such reports has been unusually small. Over 70 million people in North America use aspartame on a regular basis, yet the number of reported complaints average only around three hundred per year. The majority of complaints (67%) refer to headaches, dizziness, visual difficulties and mood alterations. Gastrointestinal problems (24%) and allergic symptoms such as hives, rashes and swelling of tissues (15%) have also been reported. On occasion seizures have been linked with aspartame exposure. In most instances these difficulties were noted only with amounts of aspartame exceeding normal use.

Double blind challenges have been carried out with aspartame. At Duke University, in one of the best designed such studies, the effects of a single large dose of aspartame in people who had claimed to be sensitive to the substance was investigated. The results showed no difference in headache frequency, blood pressure, or blood histamine concentrations (a measure of allergenic potential) between the experimental and control groups. In another study at the University of Illinois involving diabetics, subjects in the placebo group actually had more reactions than those in the aspartame group. On the other hand, surveys by physicians in headache clinics reveal that aspartame precipitates headaches about 8% of the time. This kind of conflicting data is characteristic of the research on the possible side effects of aspartame. Reported anecdotal experiences are not confirmed by carefully controlled scientific studies. This of course does not mean that the problems are not real, but does imply that in many cases the symptoms may not be caused by aspartame. People get headaches, upset stomachs, aches and pains of all kinds on a regular basis for no easily determined reason. If they recall having consumed aspartame when one of these ailments strikes, the sweetener may be judged to be guilty by association. This is even more likely if people are familiar with some of the adverse publicity that aspartame has received. At least one study has, however, confirmed allergic symptoms such as hives and swelling in sensitive individuals. It is unclear how the allergy comes about, since none of the

components of aspartame are believed to be capable of producing allergic reactions. It has been suggested that diketopiperazine, a compound which forms when aspartame decomposes may be responsible.

188

A number of theoretical possibilities have been advanced to account for aspartame associated problems. The three breakdown products of aspartame are all toxic in high doses. Phenylalanine is an essential amino acid which must be included in the diet for normal growth and maintenance, but sustained high blood levels can lead to brain damage. This is of major concern to the one out of roughly twenty thousand children who are born with an inherited condition called phenylketonurea or PKU. These children cannot metabolize phenylalanine, which then builds up to dangerous levels in their brains. The condition therefore necessitates a severe curtailment of phenylalanine intake at least for the first six years of life. This means that aspartame, due to its phenylalanine content, is not suitable for PKU sufferers and consequently requires a warning to that effect on products in which it is an ingredient.

In the general population, phenylalanine levels in the blood after aspartame ingestion are in the same range as after eating any protein containing food. Even at abusive amounts, equivalent to a child swallowing 100 sweetener tablets, levels do not rise above those which are considered to be safe in children afflicted with PKU. Dr. Richard Wurtman, a noted MIT researcher, has suggested that some of the untoward effects of aspartame may be caused by a sudden increase in brain phenylalanine levels, especially when the sweetener is consumed along with foods high in carbohydrates. Carbohydrates trigger insulin release into the bloodstream, which in turn makes it easier for phenylalanine to be absorbed by the brain. According to Wurtman, such a sudden increase in brain phenylalanine levels can cause depression, sleep problems, headaches and even seizures. These ideas have not been confirmed in human studies, and Wurtman, who uses aspartame moderately himself, feels the problem is only significant when consumption of aspartame is unnecessarily high.

The effects of aspartic acid, another aspartame breakdown product, have also been rigorously examined. Administration of extremely large amounts to non human primates produced no damage even though blood levels were greatly elevated. In humans, even high doses are quickly eliminated. Most significantly, aspartic acid levels in the blood are not increased after eating aspartame containing foods or when drinking sweetened beverages even at the rate of three drinks in four hours.

Perhaps the most unscientific accusations leveled at aspartame have involved its methanol content. It is a fact that in large doses methanol can lead to blindness and even to death. Alarmists have therefore referred to the methanol which is released from aspartame as an unsafe substance. It must be remembered, however, that there are no safe substances, only safe doses. The amount of methanol which can be released from aspartame is inconsequential in context of the overall diet. Methanol occurs naturally in foods. In fact the "natural" methanol content of fruit juice is about 2.5 times that released in the body from a diet drink using aspartame as the sweetening agent. Even at the 99th centile level of 34 mg per kg bodyweight per day, blood levels of methanol are not detectable.

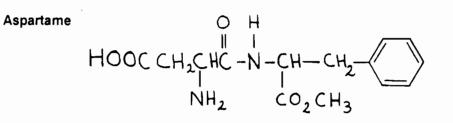
AVAILABILITY OF CYCLAMATE AND SACCHARIN

<u>C`</u>	YCLAM	ATE	SACC	HARIN
US	FA	-	FA	+
	TTS	-	ττs	+
CANADA	FA	-	FA	-
	TTS	+	TTS	*

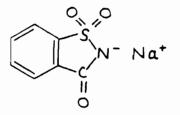
- = BANNED

+ = AVAILABLE

FA = FOOD ADDITIVE *- supermarkets; + drug stores/restaurants TTS = TABLE TOP SWEETENER



Sodium Saccharin



Sodium Cyclamate



Subject Matter Expert Comments

- What does furiously mean?

- Why would the reaction be negative?

Subject Matter Expert

- The information in this first paragraph is correct.

'- I don't think it was that strong.

- But if you eat a well-balanced diet, you shouldn't have deficiencies, and that's good for prevention of cancer or for someone who has cancer.

*- Really?

- Yes, because they eat differently when they come here

- This is the conclusion from one set of research. They

shouldn't generalize from this

- I don't think we can say that because somebody isn't

eating vegetables his risk of getting cancer is higher.

THE DIET CANCER RELATIONSHIP

"Sound nutrition is not a panacea. Good food that provides appropriate proportions of nutrients should not be regarded as a poison, a medicine, or a talisman It should be eaten and enjoyed." This statement by the Food and Nutrition Board of the National Research Council in the U.S in a 1980 publication called "Toward Healthful Diets" raised more than a few eyebrows. Reaction from consumer groups was furiously negative. These groups along with many individuals objected to the conclusion that no specific dietary advice was appropriate for all citizens. The recommendation of a balanced diet with moderation in consumption did not sit well with people who were convinced that a great many of the ills of North American society are related to improper nutrition. A document detailing the evils of food additives, the benefits of vitamin supplementation and the virtues of "organic" foods would undoubtedly have received more favorable reaction. Science however cannot deal with emotions, beliefs or anecdotal evidence", it must be based on facts stemming from well controlled and reproducible experiments. Unfortunately in the area of nutrition it/is very difficult to design and carry out studies which lead to conclusive results. Accordingly many reports of results are speckled with phrases like "may cause", "is consistent with", "is associated with", all of which imply uncertainty. The difficulty of providing "proof" one way or another in the areas of food science and nutrition leaves the door open to a variety of opinions not only among the alatmists and self styled authorities but among nutritional experts as well

Indeed, just two years after the above mentioned report the National

Research Council issued a new document entitled "Diet, Nutrition and Cancer' with more specific recommendations reflecting the state of knowledge and information pertinent to the diet and the incidence of cancer. The guidelines now recommended a reduction of fat intake from about 40% to 30% of total calories, a reduction in the consumption of cured, pickled and smoked foods and an increase in the consumption of whole grain cereal products as well as fruits and vegetables, especially those rich in carotene. Vegetables belonging to the cabbage family were highly recommended but vitamin

ntation was not advised*. The new report** was in turn

more the sugge

ion was t

e" etc. is

also criticized. Many scientists believe that not enough is known about the diet-disease connection to warrant specific guidelines for

that if the guidelines were improperty applied they could lead to nutritional deficiencies. In light of the ongoing controversy it is appropriate to examine the studies and the kind of data that have

lead to the debated recommendations. An examination of this

terms like "minerals", "vitamins", "fat, "fiber", "carote

There appears to be little doubt that many cancers are

cancer rates, suggesting an environmental influence.

controversy also serves to underline the need for a basic scientific

understanding of chemical and nutritional concepts. Familiarity with

essential* for an objective and critical discussion of the relationship

environmentally related. Epidemiological studies have clearly shown large differences in cancer rates between countries. For example, breast and colon cancer rates in many areas of the world are less than one fifth that in North America. The Japanese in turn have the

highest incidence of stomach cancer in the world. Immigrants from other countries to the U.S. and Canada however experience the local

Perhaps the best demonstration of this environmental effect comes from a study made public in 1984 by the National Cancer Research

Institute of Japan. An epidemiological study spanning 16 years and involving over 100,000 men clearly showed that the incidence of

absence of vegetables from the diet appeared to increase the risk of a wide variety of cancers. The results of the survey are summarized

cancer was greatest among those who smoked, drank alcohol, ate

mean regularly and did not consume vegetables daily. Indeed the

the population as a whole and furthe

hetween diet and cancer

below

Target Population Expert Comments

- Students will be bored by this. It's too wordy.

- There are no illustrations.

- "I'd object to that.

poor quality print pp 100

2

- 211

Target_Population_Expert

- Which National Research Council?

- "Which new report?

- Needs a reference.

- Good motivation.

- This is very heavy for first year students - they're going to get

I too fair

excited about Why not say

a typical diet at your age?

Subject Matter Expert

RELATIVE MORTALITY RATES

Smoking	Drinking	Meat	Vegetables	
No	No	No	Yes	10
No	Yes	Yes	No	11
Yes	Yes	Yes	Yes	l 7
Yes	Yes	No	No	18
Yes	No	Yes	No	18
Yes	Yes	Yes	No	25

The protective effects of vegetable consumption are dramatically illustrated by the above data; in fact even in the high risk group (smokers, dinnkers and meat eaters) the risk of cancer can be reduced by one third if vegetables are regularly eaten. This protective effect may be manifested through the fiber, Vitamin C or carotene components of the vegetables as discussed below

Accordingly many cancer experts now estimate that as much as 90% of North American cancers are environmentally determined and that a large fraction of these should therefore be avoidable "Environmental" must not be confused with "man made", in the present context the word is used to differentiate from "genetic" factors. Cigarette smoking and toxic wastes are environmental* and obviously "man made", but exposure to sunlight and the consumption of naturally occurring carcinogens can also be termed "environmental" In fact, Bruce Ames of the University of California (Berkeley) has concluded after a survey of the scientific literature that most of the carcinogens that non-smokers encounter in their daily life come from natural foods and cooking methods. For example celery and parsley contain a carcinogen which becomes activated by light, mushrooms, heans and even alfalfa sprouts contain compounds which may increase the risk of cancer. Cooking, especially when food is browned or burned adds careinogens to th dict. On the other hand, suggests Arnes, food also appears to contain natural anti-carcinopens like Vitamins C and E, selenium and carotene which may decrease the risk of the dreaded discase. The

Target Population Expert

- Tables are always troublesome.
- What are the numbers? there's no heading to that column
- "Relative" to what? There should be some explanation here.
- Students will get stuck.
- They're changing several factors here.

-I thought this was going to be a caption for the table, but it doesn't have a different typeface.

- This is getting boring.

* - I don't understand this

- This is an inappropriate level for a beginner. You really have

to tease out the information.

- Explain "natural" or don't use it, because it's misused so

much

Subject Matter Expert

- There's a danger that the public will read this as

"peanut butter causes cancer - OK I won't eat peanut butter"

- Good to allay public lears by indicating the need for

more research

fact that cancer rates aside from those related to smoking have remained almost constant over the years appears to imply that the "natural" components of the environment may be more important than the "man made" factors in inducing cancer.

In a controversial article in Science, 221,1256 (1983), Ames produced summarized the many natural foods (above) which contained various carcinogens. In this same article, he also indicated that there were many foods which were also anti-carcinogens. The main idea here was that a minimum of the questionable foods coupled with a reasonable amount of the "good" ones (vide infra) would provide as good a balance of risk/benefit as could be achieved in this very complex area. Ames was criticized by a group of 18 academics, union officials and environmentalists in a 1984 letter to Science for "Invializing" cancer risks. Ames recently published a summary of relative risk factors for cancer by a careful (but controversial) examination of the literature. The resulting index called HERP (Human Exposure dose/Rodent Potency dose). This index considers two questions: How much of the material causes considerable rates of cancer in lab animals, and how much of it might an average person be exposed to over a lifetime? The rankings do not predict a person's actual chances of developing cancer, but show comparisons. If the relative ranking of tap water is 1.0, then peanut butter (2 tablespoons:day) is 30 (aflatoxin risk) as is comfrey tea (1 cup day) (symphytine, a natural pesticide is present) (the pack day of eigeneties is rated at 12,000 while the risk of cancer from PCBs (once used in electrical transformers) is 0.2 Needless to say, such a detailed fist has created concern and discussion and will stimulate research in the future

Terget Population Expert

- There's a lot of information here. It needs a summary.

Explain that "Science" is a journal.

- This is very important, but students won't understand the

implications, so cut them out.

- What's a "questionable" lood?

- Comparisons with what? I'm getting lost.

- I'm ignoring this.

3

- This implies that not eating vegetables increases the

- True. These are all part of a balanced diet, but I

wouldn't go further than that

risk of cancer

Subject Matter Expert

- This is true.

I agree with this.

True

Subject Matter Expert

- This is a very detailed article

. This is a lunny example.

- True Good advice in general, not just for cancer

Since the second World War some 50,000 synthetic chemicals have been introduced into the environment with about 500 new ones coming into use every year. Many of these are mutagenic or carcinogenic in lab tests yet the cancer epidemic that many scientists (even Bruce Ames at one time) have predicted has not materialized. Accordingly there is widespread, though certainly not universal, belief that most cancers are caused by natural carcinogens. Many of these carcinogens are produced by plants as natural pesticides to ward off insects. Ironically the current practice of breeding insect resistant plants in order to minimize the use of synthetic pesticides may actually be introducing new carcinogens into the diet. It is also a fact of course that not everyone gets cancer even though everyone consumes natural carcinogens. The explanation for this apparent inconsistency may lie in the possibility that whereas chemicals isolated from food can cause cancer, the whole food does not. Mutagens and "anti- carcinogens" are often present in the same For example the potentially harmful effects of the psoralens food in parsicy and celery may be counteracted by the carotene and vitamin components of these foods. It appears then that attention to a scientifically balanced diet may be more important in warding off cancer than worrying about the trace amounts of synthetic carcinogens in the environment. The following sur nanzes the current state of knowledge in this important area.

The Dietary Fat-Cancer Realisionship

The above mentioned recommendation to reduce the fat content of the duct stems mustily from correlations noted by epidemiologists. A strong correlation exists between per capita fat intake and breast cancer mortality in women as well as between fat intake and mortality from colon cancer. It must be pointed out however that such associations do not imply cause. For example a similar correlation exists between gross national product and breast cancer. Although the "per capita" correlation of declary fats with cancer is strong, there appears to be no conclusive correlation of inividual fat consumption and cancer. There may be other variables in the relationship as well. Hormones like estrogen have been linked with

cancer. Does the fact that women are having fewer and later pregnancies influence the average estrogen levels ? Could it be the added calores and not the fait per se which is instrumental? The human feeding studies which would be needed to clarify the situation can never be ethically done" but studies in animals do suggest that higher levels of fait intake cause manimary tumors. Theoretically the argument can be put forward that fait cause cancer by undergoing oxidation in cells leading to the production of cancer causing reactive spocies called free radicals. These free naticals the damage the DNA of the cell, leading to improper replication. If this mechanism is correct, unsaturated faits may pose a greater risk since they are more easily oxidized. Some studies have indeed shown an association between cancer and "trans" faity acids which are produced when vegetable oils are converted into marganne. Adoquate Vitamin E. beta carotene and selenium consumption may prevent the oxidation of fais

Dr. Keith Ingold at the National Research Council in Ottawa has in fact shown that Vitamin E is the major "free-radical trapping" andoxidant in human Nood. Beta carotene can also act as an antioxidant, especially at low oxygen concentrations such as are tound in cells. It is noteworthy that this important research started out as an investigation into why engine oils break down upon exposure to oxygen in the car's engine; a nice demonstration of how important results can come from seemingly "mimportant" research Similarly the antioxidants HITT and BILA which had originally been developed to prevent faits in cereals from going rancid (and incidentially have been much maligned) may turn out to have an important role in not only the prevention of cancer but in actually solving down the eging process.

Colon cancer has also been associated with high fat, high chicksterol diets. Once again though, epidemiological studies in individuals have yielded inconsistent results. Animal feeding studies in turn have shown that dietary fat promotes colon cancer. Furthermore, populations with high rates of colon cancer have increased levels of bite acids in the feeds, these have been associated with cancer and are known to be formed in larger amounts in high fat, high cholesterol diets. In summary, the evidence may appear to be somewhat encounteratural, but the recommendation to reduce fat content by 25% does not represent a risk as long as a balanced diet is manufation.

Target Population Expert

 This is a new line of thought not related to diet. They're touching on a lot of things in an inappropriate way. I can't even evaluate the accuracy at this point. They need to define terms. This is an important point lost by being packed ina small space.

* - This should be highlighted a lot more.

- Where were the studies done?

Too dry.

- There's a need to pick the sentences apart.

- Deline estrogen.

Target ... Population Expert

*- Why not?

- Here's some relation to Chemistry. Explain free radical,

replication. This assumes a lot of pre-knowledge of such

things as oxidation, trans-, selenium.

- There's such a thing as the "Fog index"!

 For this kind of audience it needs to be more specific, less theoretical.

- What does it mean in practical terms?



.

- Good advice.

- This is controversial research - it's going to confuse the public.

- Yes, that's part of the balanced diet again.

Subject Matter Expert

- True

- This is true.

.

#- Need to define this word

The Cured Foods-Diet Association

Once again population studies have shown that cancers of the and esophagus are more common in countries such as stomac China, Japan and Iceland where the diet is high in foods that are salt cured and smoked. There is no doubt that smoke contains cancer causing compounds and salt has been reported to promote gastne cancer in rats. Sodium nitrite, a pickling agent and preservative used in cold cuts, hot dogs, ham, etc. has been linked with the potential formation of nitrosamines, known carcinogens, in the body Based upon these observations, limiting the intake of such cured or smoked foods would appear to be wise. Yet, even this dation has been challenged. It has been pointed out that recom the death rate from stomach cancer has been declining in North Amenca while the consumption of processed meats has been nsing. Furthermore, nitrite addition is so strictly regulated now that only minimal amounts are used; in fact the amount of nitrite now added can only prevent growth of the clostridium Botulinum organism if it is used in conjunction with salt. It is also true that most of the 'smoked" fouds presently marketed are smoked with liquid smoke. This is made by passing smoke through water; since the carcinogenic compounds do not dissolve in water foods "smoked" by this process are safer than "naturally" smoked toods. Although credence can be given to these enticisms, it must also be pointed out that fouds high in smoke flavor and nitrites are generally high in fat and thus in calories-perhaps enough of a reason to minimize consumption

The Selenium-Cancer Association

Scientum is a mineral required by the body in "trace" amounts. It plays a role in the activity of the enzyme glutathione peroxidase, an enzyme which protocts cells from damage by oxidation. Consistent with this activity is the observation that mammary cancer in rais fed a high polyunsaturated fat diet can be inhibited by scientum. Scientum is found in the soil and is absorbed by erops 'high soil esciplutated and differ from low soil scientum areas in several inspects' indeed lange cancer rates are lower in counting where

tobacco custans more selenium.** Mexican and Colombian tobaccos have three times as much selenium as American and British tobaccos. Some correlations between blood selenium levels and cancer have also been noted and preliminary research has shown that the selenium content of hair and nails may reflect blood levels. High intake of selenium can be toxic and the presently available information does not warrant the recommendation of supplements.

The Cancer- Vitamin C and E Connection

The evidence for this association is essentially anecdotal* although both of these vitamins are antioxidants and therefore could behave as anti-carcinogens. Vitamin E has been reported to reduce mutations in some bacternal systems and Vitamin C does block the conversion of nitrites to nitrosamines. For the latter reason Vitamin C is added to hot dogs. Similarly since both tomatoes and lettuce contain Vitamin C they can conceivably do more than just dress up the appearance and flavor of a bacon sandwich. Indeed a BLT may be the best way to consume bacon There is however no evidence that either Vitamin E or C can prevent cancer.

The Cancer-Vitamin A Connection

Remember the stories about eating carrots to see better? This may be stretching the pixel, but the vitamin A in carrota does play an essential role in the chemistry of vision. Furthermore, the vitamin and its precursor compound (beta- carotene) may also protect the bady against cancer. The rationale for this belief lites in the fact that vitamin A plays an important role in the control of cell differentiation⁶ and in that both vitamin A and especially betacarotene are efficient scavengers⁶ of chemical species called free radicals⁴⁸. Since loss of cell differentiation is a basic feature of cancerous cells and since free radicals are unstable, highly reactive chemicals which can damage our genetic materials (DNA and RNA) there is good reason to suspect that these two nutrients may have a protective effect against cancer.

Vitamin A itself can be obtained from animal products such as liver, eggs and meat or it can by ⁶ synthesized by the body from beta catoticne. Many green vegetables produce this bright orange compound but the richest sources are pumpkins, spinach and of course catotis.

Target _ Population Expert

7

- Here's a great place to talk about risk-benefit ratios

- Use practical examples. If you want to ea salami it had

better have some nitrites it otherwise it will go bad quickly and

be dangerous.

- How can they single out nitrites?

- Explain that selenium is an element.

- This could be made a lot more interesting.

'- You don't write a sentence like that in a textbook.

Target. Population. Expert

**- That's interesting.

- Very true, but I've no documentation to judge the rest.

*- I don't think students would understand the significance of

this.

- They should talk about overdoses

- Good place for Wustrations and making a strong connection

with Chemistry.

- Funny that they would talk about beliefs in such a traditional

- Explain terms marked *.

* typo.



Subject Matter Expert

- I'm worried that too much vitamin A will cause cancer

because it's a liquid soluble vitamin.

- This is only one study

The smoking is more important than the beta carolene

and cancer

Subject Matter Expert

- This is true.

. .

- That's true.

In 1975 a major epidemiological study showed that Norwegian men consuming more than the average amount of vitamin A had less than half the rate of lung cancer as compared with men having below average consumption of the vitamin. Similar findings were also reported in the following 5 years from scientists in Japan, Singapore and the United States.

A further study (Nov. 1981) published in the British medical journal Lancet supported the hypothesis that the pro-vitamin A (hela cardishe) and not the vitamin tasel was the beneficial factor. The study showed that there was an inverse relationship between intake of dietary beta-carotene and lung cancer in 1,954 middle aged male smokers over a period of 19 years. Intake of preformed vitamin A dui not show a significant effect.

Unfortunately, studies on vitamin A are often limited due to its toxicity* High levels of vitamin A lead to liver damage, headaches, lack of appetite, hair loss, menstrual problems and retarded growth in children -- problems sometimes seen among vitamin and health food faddists. On the other hand, optimal investigative approaches are possible with beta carolene since there are no known serious side effects, even with doses so high as to cause and obvious orange skin coloration. In recent years synthetic analogs of vitamin A have been prepared in an effort to reduce its toxicity. These safer compounds are now being lested with high risk groups to determine if other forms of cancer can be prevented. One such group consists of albino children in Africa who have a 100% risk of developing slan cancer. In addition, at the present time the U.S. National Institute of Health has invited all male physicians between the ages of 40 and 85 toparticipate as subjects in a placebo-controlled general study of betacarotone and cancer

A major report on this issue published in the New England Journal of Alediente, March 1984 (by the Harvard School of Public Health) explaned that although the protective effect against long cancer of beta-carotene is strongly supported by many studies, there are undications that these effects may not apply to other types of cancer

the second of those of the control that the main cause of long concerone key conservations on the tests of second other services diseases, the color control tests of the test control conservation would be the second device the control tests of the control victation. A or betation of the control tests of the control tests way.

.

The Cancer-Elber Connection Roughuge? Unappetizing, tasteless, completely indigestible but... it fishts cancer!

It all started with Dr. Dennis Burkit's 20-year observation of diets and incidence of colorectal¹⁰ cancer in rural Africa. The British surgeon noted that although cancer of the lowest five to six feet of the intestine is very prevalent in the western world it is almost nonexistent among psople in Africa consuming a high fiber diet. In Canada, about 100,000 psople get colon cancer every year, half of whom die within the same year. The same high frequency of this malignancy has been found in the U.S., Soutland, Denmark and especially New Zealand, countries which consume the highest amounts of meet and animal fat around the world.

The incidence of this type of cancer appears to be 100 times more prevalent in the lowest 1% of the small intestine. This leads scientist to believe that carcinogens are not swallowed with our food but are produced in the colon from material in the feces. It has been suggested that bie acids (biomolecules naturally released into the gut in response to the presence of fat in the diet) are chemically altened by bacteria to produce carcinogens. High colon cancer areas * have been found to be much more abundant in colorectal cancer atients than in control groups. In a recent study conducted by Dr. Tracy Wilkins, a microbiologist at the Virginia Polytechnic Institute in Blacksburg, a chemical mutagen, named fascaput nte. Was isolated from the feces of about 20 per cent of the white residents of Johanneshurg. The same compound was detected in less than 2 per cent of the rural population. The diet of the urban community is very similar to ours (high in refined carbohydrates and fat), where that of the rural population is low in mest and fat and high in fruits and vegetables. Although most carcinogens are mutagens not all mutagens are carcinogens, and therefore the presence of faccapentaene does not necessarily mean that it is the cause of cancer. Dr. David Kingston, a chemist at the Virginia Polytechnic Institute, has synthesized this compound and its cancer-causing potential will now be investigated in laboratory animals

Target Population Expert

 There's been criticism lately of studies in one population which are extrapolated to another. A present day text should point this out, perhaps in a footnote.

- Good to have explained this.

- Now I'm lost from having too much data thrown at me.

- Good.

- Difficult sentence - need to pick this apart.

- What are high risk groups?

This needs a good summary They could bring in the social

aspect by mentioning the lack of research on breast cancer

Target_ Population Expert

10

- This is the first different kind of statement - more like a

commerciai.

-Define this.

- Where are we now?

- Explain where the colon is

- This needs a picture.

* - I don't understand.

Subject Matter Expert

· They're introducing a discrepancy again.

- OK, they're resolving it here.

- It's not the traditional apple a day. It all comes down

to a balanced diet.

. This is strong. There's no such thing as an anti-cancer

menu.

That is east a balanced diel.

- This is good advice for a population, but it won't

prevent cancer in an individual

These findings certainly support the theory that fiber, which increases the rate of feces elimination, should lower one's * chances of developing cancer of the colon. However, there are some inconsistencies in the findings related to the effects of fiber. For instance. In a Canadian study** published in 1980 higher consumption of dietary fiber was shown not to have any significant effect on cancer whereas in Puerto Rico high consumption was associated with higher incidence of colon cancer. Such discrepancies may be related to the extremely heterogeneous nature of dietary liber. Dietary fiber is a mixture of indegestible chemicals: cellulose, hemicellulose, lignin and pectin.. Preliminary studies have shown that wheat bran and fiber from citrus fruits protect laboratory animals against chemically-induced colon cancer. Since citrus fruits

GUIDLENES FOR AN ANTICANCER MENU

apple a day, may not be such a bad idea.

decrease consumption of fats, nitrite-cured meats, smoked or charcoal-broiled meats and large amounts of alcohol

are also an excellent source of vitamin C (a scavenger of carcinogenic free radicals) an orange a day, or even the traditional

-increase consumption of foods rich in dietary fiber, beta carotene, vitamins $A,\, B$ and C and the mineral selenium (megadoses of dietary supplements are presently not recommended).

consume often, cruciferous vegetables such as cabhage, broccoli, Brussels sprouts and caulallower

RECENT REVIEW

A recent summary which gives a balanced report is from Scientific American, November, 1987, p. 42

Target Population Expert

- "- We can assume this means "a man's".
- There are big gender differences in this form of cancer.
- **- Who was the study done on?

- A good opportunity to talk about oat bran, which students

have heard about and which wasn't relevant.

Target Population Expert

12

General Comments

- This is Blah.

- There's too much science, too little interest. This could be

elaborated on.

- Need to know the author's training and standpoint.
- Need Illustrations.
- There's too much information without the goal of thinking who

the audience is.

- It needs more connection with Chemistry, more social

aspects, more student experiences, more references.

11

Subject Matter Expert

General Comments

- Some very precise and detailed information.
- Not enough background.
- Students need more to get an overall picture.

Subject Matter Expert's Comments

[11(M2)]

ARTIFICIAL SWEETENERS

The replacement of sugar by artificial sweeteners in the diet may be of importance in weight reduction*, in the maintenance of denial health** and in making available a greater vanety of foods for diabetics. There are three such sweeteners allowed as an additive in a large number of foods and also as a tabletop sweetener, while cyclamate (Sweet & Low) and succhann (Hernesus) are allowed only as tabletop sweeteners. In the US cyclamate has been benned since 1970, but sacchann and aspartame can but be used as additive or as tabletop sweeteners. The US. Food and Drug Administration has recently given approval to a new sweetener, acesulfame K, which could be in common use by 1989.

Sacchann, the cyclamates and aspartame were all discovered through lucky acculents. In 1879, Constantine Fahlberg, a chemist working at Johns Hopkins University in Baltimore, linked an unusual sweetness in a slice of bread he was eating to a chemical residue on his hands from the lab. Eventually the traced the sweetness to "benzoyl o-sulfomamide" which was given the name "sacchann." Fahlberg soon became wealthy but his reputation sulficred greatly when he upplied for patients in socret and refused to give any credit to his research director, fin Remsen, under whose guidance he was working when the serendiptous" discovery was made. In 1977, Nichael Sveda, working on antibacterial substances at the University of Illinois, noted a sweet taste on a cigaretic he had placed on the bench top and discovered the sweetning power of cyclamate. Similarly, aspartame was discovered in 1965 when James Schlatter, a researcher at Searle Lab mitories, licked his finger while engaged in work on anti-adcer inclucations.

Artificial sweeteners are among the most controversial food additives due to allegations of adverse health effects. These allegations include dematadopical problems, headwheat, mood variations, behavior changes, resportatory difficulties, services and cancer²⁴. A very large number of studies on these substances have been carried out with conclusions ranging from "ade under all conditions" to "unside at any dose". Scientists are elivited on the issue of artificial sweetener safety, and in many cases researchers seem to arrive at results which confirm their views. In scientific, as welf as in lay publications, supporting studies are often widely referenced while opposing results are de emphasized.

necheations

- I don't see anything on class participation or any visual equipment.

- I just read something that disagreed with this.

. "This is definitely true

- This is still quite controversial, according to the "Berkeley Wellness

Letter"

- You're always hearing this story

- 'Define this word

- ' I recently read that there is no direct link with aspartame

- There are sale doses which are easy to calculate and which might

be included here

Subject Matter Expert's Comments

- Here you have the guidelines. Make them more visible by showing the number of things that contain artificial sweeteners

- The first section is very accurate, but very dry for people with no science background. Put in some fun facts to make it easier to assimilate

- I'm having trouble digesting this

- This is good because it gives perspective to headlines such as Saccharin Causes Cancer

The general conclusion is that sweeteners when used in modest amore represent a minimal risk. Abusive amounts, which are not rare, are associated with increased, but still low risk. Most importantly, how the risk-banefit ratio[®] of artificial sweeteners is unclear. A number of ever. her of the nsk-benefit ratio[®] of artificial sweeteners is unclear. A number of stukes have shown that the use of sweeteners is not associated with weight loss Apparently, people saving on calories by taking artificial sweeteners are so proud of themselves that they compensate by indulging more extensively in other loads. The elimination of sugar from soft drinks is beneficial in terms of the dential health of children, but it would make more nutritional sense to greatly reduce the over +0 gailons per person per year rate of consumption of soft drinks altogether. Artificial sweeteners do afford real benefits to diabetica by altowing them to cal fixude which otherwise would have to be severely restricted.

SACCHARIN

Succhann is about 300 times sweeter than sugar and can be called a true "non-catoric" sweetener since it is eliminated by the body intact. It dues have a disagreeable aftertaate which can be minimized by the addition of a small amount of the amino acid, glycine, or by combining it with cyclamate. Sugar contributes not only sweetness but also "muly" to a beverage. When sugar is replaced by an antificial sweetness, but also "muly" to a sugar arabic, or carboxymethyl cellulose are often used to impart a proper multicel. as gum an mouthfeel

mouthfeel in 1972, saccharn was removed from the "Generally Recognized as Safe" hist in the U.S. when experiments showed a posible link with bladder cancer in rats. This meant that the additive was now subject to the "Delancy Clause," which had been part of an ani ndms into the Federal level, Drug and Cosnetics Act in 1958. This clause states that any additive which in any dose causes cancer in any species of animal has to be immediately banned. However, this accusatory evidence was judged to be inconclusive and saccharm use was allowed until further research either continued or related the evidence. Follow up studies on rats, mice, hamsters and monkeys were negative except for one case in which there was an increased incidence of bladder cancer in male rats fed high lifetime dores of saccharm i.e. and only if, their parents had also been fed similar high lifetime dores. These doises make up 5% of the rats diet and were the equivalent of a human consuming 800 cans of due timily ard. This of course is a very large amount, but the technique of feeding large does to rats over a short inter spins is a scientifically accepted method to estimating what may happen in humans exposed to small doses of the substance over a long pures!

Target Population Expert's Comments

- These pages are closely compacted!

- Interesting historical approach

- "Meaning of word?

Target Population Expert's Comments

2

- "How is this studied?

- Need to look at this more closely.

- "What does this mean?

- Explain these chemical names.

- The laboratory research is is an artificial comparison of anim

intake compared with humans

Subject Matter Expert's Comments

When Canadian studies in 1977 showed that the bladder cancer was indeed caused by succharin and not an impurity as some had suggested, the Canadian government banned the use of succharin as an additive. In the U.S. public and industry pressure forced Congress to pass the "Succharin Study and Labelling Act," which put a moratorium on a ban as would have been dictated by the Delaney Clause. This moratorium has been repeatedly extended. Since saccharin costs only about not twentieth as much as sugar, and since it allows for a greater variety of products and hence more shelf space, it is easy to see why the industry would oppose any restriction on the use of this high profit substance.

The cancer case against succharm is by no means iron clad. A theoretical rational for cancer induction is difficult, since succharm passes through the body unaltered. The suggestion has been made that succharm exerts its effects by temporanity attaching to and disrupting certain proteins As mentioned above, only "double generation" studies in male rats have confirmed bladder cancer, and even in these cases the incidence of tumors was seen to decline sharply with doa. The really wornsome aspects of carcinogenesis, that is tumors in several asimal species, rapid tumor formation, and tumors in various organis, have never been associated with succharm. Furthermore, human "case-control" studies have examined the infestyle and lettary intukes of a large number of bladder cancer patients without revealing any link to succharm. Neither have any studies found increased incidence of tumors antoing diabetics, a population subgroup which would be expected to be at increased risk due to extensive tuse of artificial sweeteners. It is likely then, that any bat a vanishingly small effect of eartimogen. In this is context, it should be projected to be at incidence of union badder cancer in cells which have already been exposed to some carcinogen. In this context, it should be pointed out that there are many naturally occurring substances in our dist, for example the annou acid tryptophan. Calculations based on the animal models show a "worst case scenario" of a normal diet. In all likelihood, this is also the case with succharin Calculations based on the animal models show a "worst case scenario" of a one in a million increased risk of bladder cancer due to the consumption of two saccharin sweetened beverages a day over a lifetime. Target Population Expert's Comments

- This doesn't make sense if there is an increased cancer risk.

- Needs an explanation of the study done.

-*Needs more explanation.

-Needs more background - this is pretty technical stuff!

-The meanings of some of these terms could be explained

in a sidebar.

- Good to explain the meaning of carcinogenesis to laymen.

- This paragraph seems to be justifying the US position.

- This doesn't make it seem very dangerous.

- This is fairly presented. Good to point out that Canada makes

decisions independent of the US.

Subject Matter Expert's Commente

- Again, good to explain the term "mutagenic".

- I find this very dry reading

- I'm having trouble taking it in

CYCLAMATES

Three similar compounds, namely sodium cyclamate, calcium cyclamate and cyclamic acid are collectively referred to as the "cyclamates." They are about thirty times sweeter than sugar and are chemically more stable than saccharin or aspartame. Some of a cyclohexylamine, a compound whose safety has been questioned. Like saccharin, cyclamates were granted "Generally Recognized as Safe" status with the Safe of the safety has been questioned. Like saccharin, cyclamates were granted "Generally Recognized as Safe" status with the Safe of the safety has been for a cycloantes safety has been for a cyclamate saccharin (10:1) mixture. Since saccharin was a minor component of this maxture, cyclamates were implicated in the tumor formation.

The 1969 study by the independent Food and Drug Research Laboratories was widely criticized, and never reproduced with a statistically significant number of animals. In fact, more recent studies on rats, mice, dogs, hamsters and monkeys, using both cyclamates and their breakdown product, cyclohexylamine, have shown no effect on cancer rates. Indeed, unlike sacchann, the question of whether cyclamates are carcinogenic or cocurrisogenic at all, has not been adequately answered. Neither has the allegation that cyclamates are mutagenic, that is, capable of producing genetic damage which can be inherited, been substantiated.

The only reproducible effect of cyclamates has been testicular atrophy in nats when fed large doses. This problem, which has not been seen in humans, is caused by the metabolic product cyclohexylamine. Taking into account the maximum dose at which no such effect is observed in rats, and building in a hundred fold safety factor, it is possible to arrive at an "Accepted Daily finate" of cyclamate which represents an insignificant nak. Accordingly, the Health Protection Branch in Canada allows cyclamate as a tabletop sweetner. The US hun on cyclamates is still in effect, but may be resenred in light of the fact that the Food and Drug Administration has reviewed the large number of studies submitted to it in the 1907s and has concluded that neither cyclamates nor cyclohexylamine are carcinogens. Target Population Expert's Comments

- It doesn't seem reasonable that we're not sure if it's an

effect in itself, or an effect in its compound form.

- This needs explanation - it's a whole different problem.

baked'

- I've read 200 times sweeter. Maybe this is more accurate

toods, or into components people can recognize

- It would have been interesting to show how this translates into

- But it can be added to hot foods after cooking - it's during cooking

that it breaks down. The sentence should read "can't be cooked or

ASPARTAME

Aspartame cannot be considered a "non-calone" sweetener since it is broken down in the digestive tract into its components which are absorbed and metabolized. These components, aspartic acid, phenylalanne and metabolized. These components, aspartic acid, phenylalanne and metabolized. These components, aspartic acid, phenylalanne and However, since the substance is about 180 times sweeter than sugar, very little needs to be used in foods and heverages to achieve a satisfactory degree of sweetness. Diet drinks normally contain about 60 milligrams per 100 milligrams per serving. Industry figures show that 99% of aspartame users consume less than 34 mg per kg of body weight per day. For a seventy kg person this means less than 2.38 grams. The average consumption is about 500 milligrams per day Aspartame cannot be used in cooked or baked foods since it breaks down into its components and loses its sweetening power.

Aspartame is the most widely researched food additive to have ever come on the market As with any other newly introduced substance, reports of adverse reactions were expected since no amount of lesting can preclude indowpreature reactions in a small minority of the population. In reality, the number of such reports has been unusually small. Over 70 million people in North America use aspartame on a regular basis, yet the number of reported complaints average only around three hundred per year. The majority of complaints (67%) refer to headaches, diszness, visual difficulties and moud alterations. Classitomicistical proferms (24%) and allergic symptoms such as hives, rashes and swelling of tissues (15%) have also been reported. On occasion sentures have been linked with aspartame expensare. In most instances these difficulties were noted only with amounts of aspartame exceeding normal use.

- Why are these not reported in the other lab. analyses?

Subject Matter Expert's Comments

- But is this any higher than for other foods, such as

chocolate?

· Some people will always show an allergic reaction

- Oh boy! That's some word! This is all very technical - not

something you'd see in general literature

Double blind challenges have been carried out with aspartame. At Duke University, in one of the best designed such studies, the effects of a single large dose of aspartame in people who had channel to be sensitive to the substance was investigated. The results showed no difference in headache frequency, bloud prossure, or blood histamine concentrations (a measure of allerence; noticential) behaves the errorment and volume ensure pressure, or blood histamine concentrations (a measure of allergenic potential) between the experimental and coninol group in another study at the University of litinois involving diabetic subjects in the placebo group actually had more reactions than those in the apartime group. On the other hand, surveys by physicians in headache clinics reveal that appearane precipitues headaches about 8% of the time. This kind of conflicting data is characteristic of the research on the possible side effects of appartame. Reported ancedual experimenes are nut confirmed by carefully controlled scientific studies. This of course does not more that the problem are not even by draw more their in even. mean that the problems are not real, but does imply that in many cases the symptoms may not be caused by aspartame. People get headaches, upset stormets, aches and pann of all kinds on a regular basis for no cashy determined reason if they recall having consumed aspartame when one of these allments strikes, the sweetener may be judged to be guilty by association. This is even more likely if people are tamiliar with some of the adverse publicity that aspartame when severived. At least one study has, however, confirmed allergic symptoms such as hives and swelling in sensitive individuals it is unclear how the allergy comes about, since none of the components of aspartame are believed to be capable of producing allergic reactions. It has been suggested that discomposes may be responsible mean that the problems are not real, but does imply that in many

A number of theoretical possibilities have been advanced to account for aspartame associated problems. The three breakdown products of aspartame are all toxic in high doors. Phenylatanne is an essential amono acid which must be included in the diet for normal growth and maintenance, but sustained high blond levels can lead to brain diamage. This is of major concern to the one out of toughty workly thousand children who are born with an inherited condition called phenyllectonized or PSC. These children cannot metabolize phenyllatinne, which then builds up to dangerous levels in their brains. The condition therefore necessitates a severe curtainment of phenylatinne indate at least for the first six systex of life. This means that appartame, due to its phenylatinne content, is not situable for FNC. sufferers and conseque thy requires a warming to that effect on products in consequently requires a warning to that effect on products in which it is an ingredient

Target Population Expert's Commente

- What's a "double blind challenge"?

- "Need more on this

- Need more detail on "high doses".

198

5

.

Target Population Expert's Chrimente

-Need to look at this more closely. This type of data doesn't

appear for the others

- There are too many ratios - give a range of values. Too

complicated. This will throw people off. Why not give in

terms of "per serving"? e.g.: "If you have 4 drinks...."

"If you are a 70 kg person...."

199

7

Subject Matter Expert's Comments

- And when you use aspartame as a sweetener, it's

frequently for a high-carbohydrate food.

In the general population, phenylalanine levels in the blood after aspartume ingestion are in the same range as after eating any protein containing food. Even at abusive amounts, equivalent to a child swallowing 100 sweetner tablests, levels do not nue above those which are considered to be safe in children afflicted with PKU. Dr. Richard Wurtman, a noted MIT researcher, has suggested that some of the unioward effects of aspartame may be caused by a sudden increase in brain phenylalanine levels, especially when the sweetner is consumed along with fouls high in curbohydrates. Carbohydrates trigger insulin release into the bloodstream, which in turm makes it easier for phenylalanine to be absorbed by the brain. According to Wurtman, such a sudden increase in brain phenylalanine levels. These rices have not been contirmed in human studies, and Wurtman, who uses aspartame muderately hinself, feels the problem is only significant when consumption of aspartame is unnecessantly high.

The effects of aspartic acid, another aspartame breakdown product, have also been ngorously examined. Administration of extremely large amounts to non-human primates produced no damage even though hioxid levels were greatly cleated. In humans, even high doises are quickly eliminated. Most significantly, aspartic acid levels in the blood are not increased after eating aspartame containing toxis or when draking sweetined beverages even at the rate of three drinks in four bours.

Verhaps the most unscientific accusations leveled at aspartance have involved its methanol content. It is a fact that in large doses methanol can lead to blindness and even to death. Alarmists have therefore referred to the methanol which is released from aspartame as an unsite substance. It must be remembered, however, that there are no safe substances, only safe doses. The amount of methanol which can be released from aspartame is inconsequential in context of the overall diet. Methanol occurs naturally in foods, in lact the "natural" methanol content of funt using aspartame as the sweetening agent layer at the 98th centile level of 34 mg per kg bodywight per day, blood levels of methanol are not detectable. **Target Population Expert's Comments**

- Need to explain what this means.

- Need to look at how it's used and what it's used with?

- What's the control group?

- Different research is being reported here. Is there no

primate research?

- But they need to look at build-up in the lissues too

- This is a very important point.

Subject Matter Expert's Comments

- This chart is awkward.

- You could make a chart showing the caloric value of foods

sweetened in various ways; the equivalent sweetness of

aspartame; the pros and cons of each one.

- What does FA stand for?

- Why drugstore?

- TTS information is repetitive of drugstore information.

AVAILABILITY OF CYCLAMATE AND SACCHARIN

CYCLASIATE			SACO	HARIN
US	FA		FA	•
	TTS		TTS	•
CANADA	FA		FA	
	TTS	+	TTS	•

+ = BANNED

+ = AVAILABLE

FA = FOOD ADDITIVE	 supermarkets;
	+ drug stores/restaurants

TTS = TABLE TOP SWIETENER

Tarpet Population Expert's Comments

- Separate the two charts.

- Where is the aspartame?
- When was the substance banned?
- Was there a replacement?
- Look at all possible uses.
- Look at labelling regulations.
- Make a list of the substances and where they're used -
- e.g.: gum, soft drinks, etc., maybe in a supplementary iist
- for food additives are we eating the sweetener

unknowingly?

- The "FA"s, "TTS"s don't make any sense.
- This chart is not clear at all.
- This should be clearer.
- Make Canada and US separate data.

.....

9

10

(Subject Matter Expert)

· Someone with no organic chemistry background would find

this very confusing.

- I don't see what this is going to give them

- It might be broken down as nutritive (aspartame) and non-

nutritive (saccharin).

- This is not very helpful.

иате О Н I I HOO(CH₂CHC-N-CH-CH₂-I I NH₂ CO₂CH₃

N° NO'

Sodium Cyclamate

Sodium Saccharm

- Very complex! Needs more information.

(Target Population Expert)

- What's the relationship between the substances that gives

the same sweetening effect?

- What kinds of molecules are these?

- List in order of production.

- Relate composition to breakdown.



(Subject Matter Expert) GENERAL COMMENTS :

- The bottom line should be that although these products

are useful, there is abuse, that often they're used

unnecessarily - e.g., why not use fruit juice as a sweetener?

- There are no conclusions drawn.

- it's very dry for someone with no chemistry background.

- You can make learning easier by making it fun. Give the

same information, but make it lighter.

- You could put in some humour - e.g., a cartoon of fat lady

in cateleria line with Black Forest cake and a diet soft drink

(Terget. . Population . Expert)

GENERAL COMMENTS :

- There's a lot of technical stuff here.

 It could read as a whodunnit - e.g. scenario of history, use, research and ban on each one -- what kind of research was done and what motivated it?

- It's quite wordy.

 It could be made a little more historically based and spiced up, e.g.: "Tracking down the hidden killer in saccharin".

- Was the research rigorous? Or can we accept the risk?

- Does II warrant further research? Lab research doesn't simulate the risk population in humans.

- it needs background on allergies and what the FDA looks for.

 It needs a pronunciation guide for the chemical names to help students aboots them - or put them into layman's terms.

 It's not a stanted document, but what conclusion is the learner to draw? that the research is not rigorous? too rigorous?

- Say at the end: "Nothing could have survived this thoroughness".

[12(M1)]

THE DIET CANCER RELATIONSHIP

PostTest

(Qu 10) 82%

PostTest

PreTest

PreTest

"Sound nutrition is not a panacea. Good food that provides appropriate proportions of nutrients should not be regarded as a poison, a medicine, or a talisman It should be eaten and enjoyed * This statement by the Food and Nutrition Board of the National Research Council in the U.S in a 1980 publication called "Toward Healthful Diets" raised more than a few eyebrows. Reaction from consumer groups was funously negative. These groups along with many individuals objected to the conclusion that no specific dietary advice was appropriate for all citizens. The recommendation of a halanced diet with moderation in consumption did not sit well with people who were convinced that a great many of the ills of North American society are related to improper nutrition. A document detailing the evils of food additives, the benefits of vitamin supplementation and the virtues of "organic" foods would undoubtedly have received more favorable reaction. Science however cannot deal with emotions, beliefs or anecdotal evidence, it must be based on facts stemming from well controlled and reproducible experiments. Unfortunately in the area of nutrition it is very difficult to design and carry out studies which lead to conclusive results. Accordingly many reports of results are speckled with phrases like "may cause", "is consistent with", "is associated with", all of which imply uncertainty. The difficulty of providing "proof" one way or another in the areas of food science and nutrition leaves the door open to a variety of opinions not only among the alarmists and self styled authorities but among nutritional experts as well

Learner Comments

- Define panacea (3).

Learner Comments

2

indeed, just two years after the above mentioned report the National Research Council issued a new document entitled "Diet, Nutrition and Cancer* with more specific recommendations reflecting the state of knowledge and information pertinent to the diet and the incidence (Qu 1) 26% (Qu 1) 88% of cancer. The guidelines now recommended a reduction of fat intake from about 40% to 30% of total calories, a reduction in the consumption of cured, pickled and smoked foods and an increase in the consumption of whole grain cereal products as well as fruits and vegetables, especially those rich in carotene. Vegetables belonging to the cabhage family were highly recommended but vitamin supplementation was not advised. The new report was in turn also (Qu 11) 79% criticized. Many scientists believe that not enough is known about the dict-disease connection to warrant specific guidelines for the population as a whole and furthermore the suggestion was made that (Qu 12) 68% it the guidelines were improperty applied they could lead to nutritional deficiencies. In light of the ongoing controversy it is appropriate to examine the studies and the kind of data that have lead to the debated recommendations. An examination of this controversy also serves to underline the need for a basic scientific understanding of chemical and nutritional concepts. Familianty with - Is carotene a nutrient? terms like "minerals", "vitamins", "fat, "fiber", "carotene" etc. is essential for an objective and critical discussion of the relationship - Define the terms (2) between diet and cancer *- Define (2). There appears to be little doubt that many cancers are environmentally related. Epidemiological# studies have clearly shown large differences in cancer rates between countries. For example, breast and colon cancer rates in many areas of the world are less than one fifth that in North America. The Japanese in turn have the highest incidence of siomach cancer in the world. Immigrants from other countries to the U.S. and Canada however experience the Explain why. local cancer rates, suggesting an environmental influence Perhaps the best demonstration of this environmental effect comes from a study made public in 1984 by the National Cancer Research Institute of Japan. An epidemiological study spanning 16 years and in olving over 101/00) men clearly showed that the mendence of cancer was greatest among those who smoked, drank alcohol, are ment regularity and did not consume vegetables daily. Endeed the atorscent organized to the distrappened to increase the risk of a when are to be another. The results of the savey are summarized (Qu 2) 85% (Qu 2) 94°.

)						
PreTest	PostTest					Learner Comments
		RELATIVE MORTALI	Y RAILS			3
		Smoking Drinki	ng Meat	Vegetables		- Reorganize by topic; or by yes's then no's.
		No No	No	Yes	1.0	- Chart is not clear (2).
	(012) 448	No Ye	y Yes	No	11	- Explain the numbers (3).
	(Qu 13) 44%	Yes Yes	Yes	Yes	17	- Boldface the differences.
		Yes Yes	No	No	18	- There's missing information on the smokers re, their sleeping habits, type
		Yes No	Yes	No	18	cigarette, working environment, exercise etc.
		Yes Yes	Yes	No	2.5	
1		The protective effects of 0 illustrated by the above d (smokers, dinikers and m by one third if vegetables nai) be manifested throug components of the vegeta	lata; in fact eve eat eaters) the i are regularly o th the filter, Vi	en in the high r risk of cancer e caten. This prot tamin C or caro	nsk group an he reduced ective effect	
(Qu 3) 97%	(Qu 3) 85%	Accordingly many cancer of North American cancer large fraction of these shi	s are environm	nentally determs		*Accordingly*?
		"Law ronmental" must no present context the word factors: Cigarette smoking obviously "main made", h consumption of naturally "environmental". In fact, (Berkeley) has concluded that most of the careings	is used to diffe g and toxic was ut exposure to occurring care Bruce Arnes o after a survey	erentiate from " stes are environ sunlight and th mogens can also if the University of the scientifi	genetic" mental and ic o be termed o California c literature	- Give examples of "environmental" and "man made"
	(Qu 14) 94°。	daily life come from natu example celery and parsle activated by light mushin contain compounds which especially when listed is b diet on the other hand is indictation of calencies la catolica which may deer	ral foods and c y contain a car norms, beans an i may increase covined or bui aggests. Vines, ke Vitannins C	cooking method reinogen which id even alfalla s the risk of cane ned adds carein dock also appe l'and l. setenin	s For becomes sprouts cer Cooking, ogens to the ars to contain m and	
PreTest	PostTest	fact that cancer rates aside remained almost constant o "natural" components of th	wer the years a e environment	appears to imply may be more in	that the	Laarner Comments 4
		than the "man made" factor in a controversial article in produced summarized the r contained various careinogy that there were many foods main idea here was that a n excepted with a reasonable would provide as good a b in this very complex area. A academics, union officials a Science for "invisitizing" or summary of relative risk far controversial) examination	Science, 221, many natural for many natural for many natural for amount of the amount of the amount of the amount of the amount of the amount of the same of risk/ A mes was critis and environme uncer risks Am cors for cancer of the literatur	1256 (1983), Ai soods (above) wh ne article, he als ao anti-carcinog e questionable f "good" ones (v emefit as could cized by a grou mealists in a 191 nes recordly put re. The resulting re. The resulting	nch so indicated pens. The oods side infra}" be achieved p of 18 54 letter to blished a sut g index.	- Awkward wording. *- Explain meaning.
	(Qu 15) 79%	called IIERP (Human Exp index considers two questi considerable rates of cancer might an average person be do not predict a person's ac show companisons. If the ri- peanut butter (2 tablesprov as is comfrey tea (1 cup da	ns: How much r in tab animals exposed to ov tual chances of clative ranking ns.day) is 30 (i	h of the materia s, and how muc ver a lifetime? T f developing ca of tap water is aflatoxin risk)	I ^{ss} causes h of it he rankings neer, but 1.0, then	**- What material? - What do the numbers mean? - Explain "allatoxin" (2)

peant nutre (2 tantespions day) is 50 (aniaoan risk) as is comfrey tea (1 cup day) (symphytine, a natural pesticide is present) One puck day of ergarettes is rated at 12,000 while the risk of cancer from PCBs (once used in electrical transformers) is 0.2 Needless to say, such a detailed list has created concern and discussion and will stimulate research in the luture.

PreTest	PostTest		203 Learner Commente s
(Qu 4) 79%	(Qu 16) 91% (Qu 4) 94%	Since the second World War some 50,000 synthetic chemicals have been introduced into the environment with about 500 new ones coming into use every year. Many of these are motagenic ⁶ or carcinogenic in lab tests yet the cancer epidemic that many scientists (even Bruce Arnes at one time) have predicted has not materialized Accordingly there is widespread, though certainly not universal, heliet that most cancers are caused by natural carcinogens. Many of these carcinogens are produced by plants as natural pesticides to ward off insects. Ironically the current practice of breeding insect resistant plants in order to minimize the use of synthetic pesticides may actually be introducing new carcinogens into the diet. It is also a fact of couste that not everyone gets cancer even though everyone consumes natural carcinogens. The explanation for this apparent inconsistency may lie in the possibility that whereas chemicals isolated from tool can cause cancer, the whole food does not. Mutagens and "anni-carcinogens" are often present in the same food for example the potentially harmful effects of the poralens in paraley and cellery may be counteracted by the carcene and vitamin components of these foods. It appears the materiatent on a scientifically balanced deet may be more important in warding off cancer than worrying about the trace amounts of synthetic carcinogens in the environment. The tollowing summarizes the current state of knowledge in this important area	•• Explain
(Qu 5) 65%₀	(Qu 5) 53%	The Dietary Fat-Cancer Realationship The above mentioned recommendation to reduce the fat content of the diet stems mostly from correlations noted by epidemiologists A strong correlation exists between per capita fat intake and breast cancer mortality in women as well as between fat intake and mortality from colon cancer. It must be pointed out however that such associations do not imply cause. For example a similar correlation exists between gross national product and breast cancer. Although the "per capital" correlation of dietary fats with cancer is strong, there appears to be no conclusive correlation of individual fat consumption and cancer * There may be other variables in the relationship as well. Hormones like estrogen have, been linked with	*- This sentence should come before "For example * - I don't understand the sentence.

PreTest	PostTest		Learner Comments 💪
		cancer. Does the fact that women are having fewer and later	
		pregnancies influence the average estrogen levels ? Could it be the	
		added calories and not the fat per se which is instrumental? The	
		human leeding studies which would be needed to clarify the	
		situation can never he ethically done but studies in animals do	
		suggest that higher levels of fat intake cause mammary tumors.	 Give an example of tree radicals.
	(Qu 17) 91%	Theoretically the argument can be put forward that fats cause cancer	This is the second it second more surface to a
	(00 17) 3176	by undergoing oxidation in cells leading to the production of cancer causing reactive species called free radicals. These free radicals then	 This is too concise, it needs more explanation.
		damage the DNA of the cell, leading to improper replication If this	- I thought saturated fats were worse
		mechanism is correct, unsaturated fats may pose a greater risk since	- I ubugin saturated iets were worse
	(Qu 18) 71%	they are more easily oxidized. Some studies have indeed shown an	
	(11)	association between cancer and "trans"s fatty acids which are	
		produced when vegetable oils are converted into margarine	*- Explain
	(Qu 19) 35%	Adequate Vitamin E., beta carotene and selenium consumption may	
		prevent the oxidation of fats.	- Difficult to understand.
(O., C) (AAR()	(O	Dr. Keith Ingold at the National Research Council in Ottawa has in	
(Qu 6) (44%)	(Qu 6) 79%	fact shown that Vitamin E is the major "free-radical trapping" anti-	
		oxidant in human blood. Beta carotene can also act as an	
		antioxidant, especially at low oxygen concentrations such as are	
		found in cells. It is noteworthy that this important research started out as an investigation into why engine oils break down upon	
		exposure to exygen in the car's engine; a nice demonstration of how	
		important results can come from seemingly "unimportant" research.	
		Similarly the antioxidants BEFF and BEEA which had originally	
		been developed to prevent fats in cereals from going rancid (and	
		incidentally have been much maligned) may turn out to have an	- Should be more concise.
		important role in not only the prevention of cancer but in actually	
		slowing down the aging process	-Clear up the contradictions.
(O - 7) 7 (A)	10. 3. 30		
(Qu 7) 74°6	(Qu 7) 79%	Colon cancer has also been associated with high fail, high	
		cholesterol diets. Once again though, epidemiological studies in	
		individuals have yielded inconsistent results. Animal feeding	
		studies in turn have shown that dietary fat promotes colon cancer.	
		Furthermore, populations with high rates of colon cancer have increased levels of bile acids in the feees, these have been associated	
		with cancer and are known to be formed in larger amounts in high	
		tat, high choicsterol diets. In summary, the evidence may appear to	
		be somewhat circumstantial*, but the recommendation to reduce lat	* meaning?
		content by 25% does not represent a risk as long as a balanced diet	
		is manifanted	

PostTest Learner Comments 7 PreTest 204 The Cured Foods Diet Association (Qu 20) 88% Once again population studies have shown that cancers of the stomach and esophagus are more common in countries such as China, Japan and Iceland where the diet is high in foods that are sail cured and smoked. There is no doubt that smoke contains cancer causing compounds and salt has been reported to promote gastric cancel in rais. Sodium nitrite, a pickling agent and preservative used in cold cuts, hot dogs, ham, etc. has been linked with the potential *- Define formation of nitrosamines*, known carcinogens, in the body. Based upon these observations, limiting the intake of such cured or smoked toods would appear to be wise. Yet, even this recommendation has been challenged. It has been pointed out that the death rate from stomach cancer has been declining in North America while the consumption of processed meats has been using Furthermore, nitrite addition is so strictly regulated now that only minimal amounts are used, in fact the amount of nitrite now added can only prevent growth of the clostridium Hotulinum organism if it is used in conjunction with salt. It is also true that most of the 'smoked" foods presently marketed are smoked with liquid smoke This is made by passing smoke through water, since the careinogenic compounds do not dissolve in water foods "smoked" · Unclear sentence. by this process are safer than "naturally" smoked toods. Although credence can be given to these criticisms, it must also be pointed out that foods high in smoke flavor and nitrites are generally high in fat and thus in calories-perhaps enough of a reason to minimize consumption The Scientum Cancer Association Selenium is a nimeral required by the body in "trace" amounts. It (Qu 21) 76% plays a role in the activity of the enzyme glutathione peroxidase, an Define selenium enzyme which protects cells from damage by oxidation. Consistent with this activity is the observation that mammary cancer in rais fed a high polyansaturated fat diet can be inhibited by scienium Sclemum is found in the soil and is absorbed by crops. High soil sclemum areas correlate inversely with cancer but these areas are also less populated and differ from low soil scienting areas in several respects didect funge incertales are lower in countries where PreTest PostTest Learner Comments Q tobacco contains more scientum. Mexican and Colombian tobaccos have three times as much scientium as American and British tobaccos. Some correlations between blood selenium levels and cancer have also been noted and preliminary research has shown that the scientum content of hair and nails may reflect blood levels. High intake of scienium can be toxic and the presently available mendation of supplements information does not warrant the recom-The Cancer- Vitamin C and E Connection The evidence for this association is essentially anecdotal although both of these vitamins are antioxidants and therefore could behave

as anti- carcinogens. Vitamin E has been reported to reduce mutations in some bacterial systems and Vitamin C does block the conversion of nitrites to nitrosamines. For the latter reason Vitamin C is added to hot dogs. Similarly since both tomatoes and lettuce contain Vitamin C they can conceivably do more than just dress up the appearance and flavor of a bacon sandwich. Indeed a BLT may (Qu 8) 47% (Qu 8) 47% he the best way to consume bacon. There is however no evidence that either Vitamin Is or C can prevent cancer. The Cancer-Vitamin A Connection Remember the stories about eating carrots to see better? This may be stretching the point, but the vitamin A in carrots does play an essential role in the chemistry of vision. Furthermore, the vitamir (Qu 22) 88*; and its precursor compound* (heta- carotene) may also protect the * Explain term. body against cancer. The rationale for this belief lies in the fact that vitamin A plays an important role in the control of cell differentiation and in that both vitamin A and especially betacarolene are efficient seavengers of chemical species called free radicals. Since loss of cell differentiation is a basic feature of - Need to explain the chemistry of this. cancerous cells and since free radicals are unstable, highly reactive chemicals which can damage our genetic materials (DNA and RNA). · Don't understand the last sentence there is good reason to suspect that these two nutrients may have a protective effect against cancer Vitanin Vitself can be obtained from animal products such as liver, eggs and must or it can be synthesized by the bady from thetacarotene. Many green cegetables produce this bright orange compound but the richest sources are pumpkins, spinach and of and and

	PreTest	PostTest		Learner Comments 205	9
0	(Qu 9) 38%	(Qu 9) 85%	In 1975 a major epidemiological study showed that Norwegian men- consuming more than the average amount of vitamin A had less than half the rate of lung cancer as compared with men having below average consumption of the vitamin. Similar findings were also reported in the following 5 years from scientists in Japan, Singapore and the 1 micd States. A further study (Nov. 1981) published in the British medical journal Lancet supported the hypothesis that the pro-vitamin A theta carotene) and not the vitamin (tself was the beneficial factor The study showed that there was an inverse relationship between intake of dictary beta-carotene and lung cancer in 1.954 modile aged male studies was a period of 19 years. Intake of preformed vitamin		
			A did not show a significant effect. Unfortunately, studies on vitamin A are often limited due to its toxicity. High levels of vitamin A lead to liver damage, headaches, lack of appetite, hair loss, mensirial problems and retarded growth in children problems sometimes seen among vitamin and health fool Laddists. On the other hand, optimal investigative approaches are possible with beta carotene since there are no known serious side effects, even with doses so high as to cause and obvious orange skin coloration*. In recent years syndictic analogs of vitamin A have been prepared in an effort to reduce its toxicity. These safer	- Meaning of the first sentence? • Are they serious?	
			compounds are now being tested with high risk groups to determine if other terms of cancer can be prevented. One such group consists of albino children in Atrica who have a 100% risk of developing skin cancer. In addition, at the present time the U.S. National Institute of Health has invited all male physicians between the ages of 40 and K5 to participate as subjects in a placeho-controlled general study of beta-carotene and causer.	- I assume they mean albino black children. - Needs to be clarified.	
			A major report on this issue published in the New England Journal of Medicine, March 1984 (by the Harvard School of Public Health) explained that although the protective effect against lung cancer of beta carotene is strongly supported by many studies, there are indications that these effects may not apply to other types of cancer		
			In conclusion, it should be noted that the main cause of lung cancer, smoking, also increases one strisk of several other serious diseases, including atheroselerosis ¹⁰ – a primary cause of death in North America. However, there is no evidence that either vitamin A or hela carotine, affects this condition in any way.	"- Deline (2) - "However" doesn't make sense here.	
	PreTest	PostTest		Leerner Comments	
			The Cancer-Fiber Connection		10
		(Q 23)100%	Roughage? Unappetizing, tasteless, completely indigestible but it lights cancer!	- If they're going to bring in a more intimate tone, it should have been	•
				embedded throughout the whole work.	
			It all started with Dr. Dennis Burkitt's 20-year observation of dicts and incidence of colorectal cancer in rural Africa. The British		
			surgeon noted that although cancer of the lowest five to six feel of the intestine is very prevalent in the western world it is almost	- Does it mean that Africans in general eat a high fiber diet, or that an	
			ownewstent among people in Africa consuming a high fiber duet. In Canada, about 100,000 people get colon cancer every year*, hall of whom de within the same year. The same high frequency of this	experimental group was ted a high fiber diet? - How old are they? Is it important to know that?	
			malignancy has been found in the U S, Scotland, Denmark'and especially New Zealand, countries which consume the highest amounts of meat and animal fat around the world		
			especially New Zealand, countries which consume the highest	**- Explain what this is.	

11

12

PreTest

PostTest

These findings certainly support the theory that fiber, which increases the rate of feees elimination, should lower one's chances of developing cancer of the colon. However, there are some inconsistencies in the findings related to the effects of fiber. For instance, in a Canadian study published in 1980 higher consumption of dietary fiber was shown not to have any significant effect on cancer whereas in Puerto Rico high consumption was associated with higher incidence of colon cancer. Such discrepancies may be related to the extremely heterogeneous nature of dietary fiber Detary fiber is a mixture of indigestible chemicals, cellulose, hemicellulose, lignin and pecun – Preliminary studies have shown that wheat bran and liber from citrus fruits protect laboratory animals against chemically-induced colon cancer. Since citrus fruits are also an excellent source of vitamin C (a scavenger of carcinogenic free radicals) an orange a day, or even the traditional apple a day, may not be such a bad idea.

GUDELINES FOR AN ANTICANCER MENU

-decrease consumption of fats, nitrite-cured meats, smoked or charcoal-broiled meats and large amounts of alcohol

increase consumption of foods nch in dietary fiber, beta carotene, vitamins A, I; and C and the mineral selenium (megadoses of dietary supplements are presently not recommended)

consume often, cruciterous vegetables such as cabbage, broccoli, Brussels spronts and cauliflower

RECENT REVIEW

A recent summary which gives a balanced report is from Scientific American, November, 1987, p. 42

Learner Comments

- Did they deal with alcohol?

- Needs examples
- What foods contain selenium?

Learner Comments

GENERAL COMMENTS OF LEARNERS

- Not enough commas.
- Needs a summary at the end.
- Needs summaries to help studying.
- Needs keywords like "on a side note..." or "most importantly..."
- Needs sidebars.
- You have to pay a lot of attention to "may", "might", and "could" in this.
- There's a problem with tone.
- It would be more interesting to see the science behind it instead of the

results of the studies.

- Should be broken up more.

PreTest

PostTest

		[12 M2)]	
PreTest	Post Test	ARTIFICIAL SWEETENERS	Learner Comments
		The replacement of sugar by artificial sweeteners in the diet may be of importance in weight reduction, in the maintenance of dental health and in	- Had to read the paragraph twice
	(Qu 9) 76%	making available a greater variety of foods for diabelies. There are three such swecteners allowed in Canada today. Aspartame (trade names Equal	- Didn't know there were so many different sweeteners - I thought
	(Qu 10) 47%	and Nutratweet) is allowed as an additive in a large number of foods and also as a tabletop sweetener, while exclamate (Sweet & Low) and saccharm (Hermestas) are allowed only as tabletop sweeteners. In the U.S. eyclamate	they were just different trade names.
		there is a subject of the state	- If they're so dangerous, why are they offered at all?
		K, which could be in common use by 1989.	- This is all Greek to me.
(Qu 1) 59°₀	(Qu 1) 100%	Socchain, the Cyclamates and aspartame were all discovered through locky accidants. In 1879: Constantine Fahlberg, a chemist working al Johns Hopkins it uversity in Baltimore, Inked an unusual sweetness in a site of bread he was cating to a chemical residue on his hands from the lab Eventually he traced the sweetness to "benzoy" ossificiantimade which was given the name "socchain" Fahlberg soon became wealthy but his reputation suffered greatly when he applied for patents in secret and refused to give any credit to his research director. If a Remsen, under whose guidance he was working when the seendaptious discovery was made. In 1947, Micheel Sveed, working on antitracterial substances? at the	 Confusing and irrelevant. Is it put in for novelty value? This type of information isn't given for the other sweeteners, so why this one?
		University of Illinois, noted a sweet taste on a creatette he had placed on the bench top and discovered the sweetening power of cyclamate. Similarly, aspartance was discovered in 1965 when James Schlatter, a researcher at Scarle Laboratores, licked his tinger** while engaged in work on anti- uleer inclications.	* Is this relevant? ** Yuk! - This doesn't sound very scientific!
	(Qu 12) 76%	Artificial sweet nets are among the most controversial food additives due to allegations of adverse health effects. These allegations include dermatological problems, headaches, mood variations, behavior changes, respiratory difficulties services and cancer. A very large number of studies on this is substances have been carried out with conclusions ranging from "rade under all conditions" to "mostle at any doie." Scientistics are divided on the issue of artificial sweetenet solety, and in many cases researchers seem to arrive at results which continuit ther research areas in seema and a	 How can there be this range? There have to be some criteria or guidelines.
		in Lay public autors, supporting studies are often widely referenced while opposing results are de employized.	- What views? Not clear.
			- is this para needed? the fourth para, says the same thing better

 \bigcirc

(PreTest	PostTest)		(Learner Gomments)
		The general conclusion is that sweeteners when used in modest amounts represent a minimal risk. Abusive amounts, which are not rare, are associated with increased, but still low risk. Most importantly, however, the n-k-benefit ratio* of artificial sweeteners is unclear. A number of	- Repetitive. Skip the tirst sentence "which are not rare" doesn't make sense.
		studies have shown that the use of sweeteners is not associated with weight loss. Apparently, people saving on calones by using artificial sweeteners are so provid of themselves that they compensate by indulging	- This is shifty. It would have been a lot easier to have said that the
	(Qu 13) 41	more extensively in other foods. The elimination of sugar from soft drinks is beneficial in terms of the dental health of children, but it would make	range of findings was varied and that the information is inconclusive
		more nutritional sense to greatly reduce the over 40 gallons per person per year rate of consumption of solt drinks altogether. Artificial sweeteners do altout real benefits to diabetics by allowing them to eat foods which	- "Then why do we have artificial sweeteners?
	(Qu 11) 71%	where is would have to be severely restricted.	- There's a change of lone here. This isn't science any more
		SACCHARIN	
	(Qu 14) 35%	Sacchann is about 300 times sweeter than sugar and can be called a true "non-calone" sweetener since it is eliminated by the body intact.* It does have a disagreeable altertaste which can be minimized by the addition of a	- * Meaning?
		small amount of the amono acid, glycine, or by combining it with cyclamate. Sugar contributes not only sweetness but also "body" to a beverage. When sugar is replaced by an artificial sweetner, thickeners such	- But cyclamate has been banned since 1970. It doesn't say why it
	(Qu 15) 59%	as gum arabic, or carboxymethyl cellulose are often used to impart a proper mouthleet	was banned.
		the second se	- Are we to assume that the thickeners are sale?
(Qu 2) 100%	(Qu 2) 41%	In 1972, succhain was removed from the "Generally Recognized as Safe" list in the US when experiments showed a possible link with bladder cancer in rats. This meant that the additive was now subject to the "Delaney Clause," which had been part of an amendment to the Federal Food, Frig and Cosnetics Act in 1958. This chaise states that any additive which in any dose causes cancer in any species of animal has to be immediately banned. However, the accusatory evidence was pulged to be inconclusive and succharm use was allowed until further research either continued or retuted the evidence. Follow up studies on rats, mee,	- This is a good paragraph; it makes sense
		hamsters and monkeys were negative except for one case in which there was an increased incidence of bladder cancer in male rats fed high lifetime	- Strange that it's only males. Did they study females? (twice)
		doses of saccharn it, and only if, their parents had also been led similar high lifetime doses. These doses made up 5% of the rats diet and were the	"High lifetime doses" is not clear.
		equivalent of a human consuming 800 cans of diet drink per day. This of course is a very large amount, but the technique of feeding large doses to tats over a short time span is a scientifically accepted method of estimating	-If our parents are saccharm, will our children have bladder cancer?
		what may happen in humans exposed to small doses of the substance over a long period	(fwice)

(PreTest	PostTest)		(Learner Comments)
(Qu 3) 59%	(Qu 3) 94% (Qu 16) 76%	When Canadian studies in 1977 showed that the bladder cancer was indeed caused by succhann and not an impurity as some had suggested, the Canadian government banned the use of sacchann as an additive. In the U.S. public and industry pressure forced Congress to pass the "Succhann Study and Labelling Act," which put a moratorium on a ban as would have been dictated by the Delaney Clause. This moratorium has been repeatedly extended. Since sacchann costs only about one (wentieth as much as sugar, and since it allows for a greater variety of products and hence more shell study: it is easy to see why the industry would oppise any restriction on	- Does this mean they banned it then put it back on the market?
		space, it is easy to see will de mousing would oppree any resolucion on the use of this high profit substance	-' meaning? How does this relate?
		The cancer case against succhann is by no means iron clad. A theoretical rational for cancer induction is difficult, since succharin passes through the	- This para. Is O.K. for somebody in medicine - not clear to me.
		body smaltered. The suggestion has been made that saccharm events its effects by temporarily attaching to and disrupting certain proteins. As mentioned above, only "double generation" studies in male rats have	- Not expressed well (twice).
		contirmed bladder cancer, and even in these cases the incidence of tumors was seen to decline sharply with dose. The really worrisome aspects of	- If it comes out with everything it goes in with, what does it leave
		carcinogenesis, that is tumors in several animal species, rapid tumor formation, and tumors in various organs, have never been associated with succhain forthermore, human "case-counter" studies have examined the	behind that causes cancer?
		Intestyle and dietary intakes of a large number of bladder cancer patients without revealing any link to saccharin. Neither have any studies found	- I don't know enough Chemistry for this (twice).
		mercased mendence of tuniors among diabetics, a population subgroup which would be expected to be at increased risk due to extensive use of artificial sweetners. It is hiely then, that any but a wanshingly small	- Awkward sentences. I had to read this three times. You expect
		effect of saechann on human bladder cancer would have been detected. The possibility exists that saecharin is a "co-carcinogen,"# that is a substance	this to be related to saccharin, then the end of each sentence
		which may initiate cancer in cells which have already been exposed to some care incore in this context of should be pointed out that there are many naturally occurring softstances in our diet, for example the amino acid	proves that it's not.
		tryptophan, which are known to promote cancer at high doses but represent no risk at all as part of a normal diet. In all likelihood, this is also the case	- What's a carcinogen? co-carcinogen? carcinogenesis? Explain
		with sacchaim cladulations based on the available animal models show a "worst case security" of a one in a million increased risk of bladder cancer due to the consumption of two sacchain sweetened beergages a day over a	these first before explaining co-carcinogen
		life time.	 This is hard if you don't know any Chemistry.

(Qu 17) 71%

PostTest) (PreTest (Learner Commente) CYCLAMATES Three similar compounds, namely sodium cyclamate, calcium cyclamate and cyclamic acid are collectively referred to as the "cyclamates." They are about thirty times sweeter than sugar and are chemically more stable than sacchann or aspartame. Some of a cyclamate dose is excreted by the body unchanged, but some is converted to a cyclohexylamine, a compound whose safety has been questioned." Like saccharin, cyclamates were granted "Ciencrally. Recognized as Safe" status was resended in 1969 when bladder tumors were discovered in nata which had been 1969 when bladder tumors were discovered in nata which had been to a cyclamate-saccharin (10.1) mixture. Since saccharin was a muor component of this nuxture, cyclamates were implicated in the tumor formation. (Qu 4) 41% (Qu 4) 94% Needs more detail -Were the tumors cancerous? The 1969 study by the independent Food and Drug Research Laboratories was widely enticized, and never reproduced with a statistically significant number of animals in fact, more recent studies on rats, mice, dogs, hamslers and inonkeys, using both epclanates and their breakdown product, cyclohevylanine, have shown no effect on cancer rates. Indeed, unlike^a sacchain, the question of whether cyclimates are carcinogene or co-carcinogenic at all, has not been adequately answered. Neither has the allegation that cyclimates are usinghere, that is, capable of producing genetic damage which can be inherited, been substantiated. (Qu 5) 59% (Qu 5) 65% - "It's not answered for saccharin either! - Define mutagenic better. I can figure out the meaning from the context, but it should be defined separately. - What relevance does rat research have to humans?

The only reproducible effect of cyclamates has been testicular attophy in rate when ted large dones. This problem, which has not been seen in humans is caused by the metabolic product cyclohexylamine. Taking into account the nearminum does at which no such effect is observed in rates, and building in a hundred fold safety factor, it is possible to arrive at an "Accepted Daily Intake" of cyclamate which represents an insignificant risk. Accepted Daily Intake" of cyclamate which represents an insignificant risk. Accepted I and the test on Branch in Canada allows cyclamate as a tabletop sweetener. The U.S. ban on cyclamates is still in effect, but may be resented in fight of the fact that the Food and Drug Administration has reviewed the large number of studies submitted to it in the DMSs and has concluded that neither cyclamates nor cyclohexylamine are calcingetis. are caternogens

- Canada/US differences should be made clearer.

- Contradicts what was said in the last paragraph . has it been

answered or not?

- These words are longue-twisters.

			209
(PreTest	PostTest)	ASPARTAME	(Learner Gomments)
0	(Qu 19) 88%	Aspartame cannot be considered a "non-calone" sweetener since it is broken down in the digestive tract into its components which are absorbed and metabolized. These components, aspartic acid, phenylalamme and methanol account for the 4 calones per gram energy rating of aspartame. However, since the substance is about 180 times sweeter than sugar, very http://doi.org/10.1003/since.com/since.com/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/since/	- This explains non-caloric sweetener better. They should have said
			this before.
	(Qu 18) 71%		 I understand the words in the first sentences, but not the meaning.
		Industry figures show that 99% of aspartame users consume less than 34 ing per kg of tody weight per day. For a seventy kg person this means less than 2 kg grams. The average consumption is about 500 milligrams per day.	- There's a lot of information in this para. It needs to be read twice
		Aspartance cannot be used in cooked or baked foods since it breaks down into its components and loses its sweetening power	 So what's sale?
(Qu 7)	29% (Qu 7) 76%	Aspartume is the most widely researched food additive to have ever come	
		on the market. As with any other newly introduced substance, reports of adverse reactions were expected since no amount of testing can preclude idios relate reactions in a small minority of the population. In reality, the	- Here again, they build it up, then the bottom line is that they can't
		number of such reports has been unusually small Over 70 million people in North America use asparation on a regular basis, yet the number of reported complaints average only around three bundred per year. The	prove anything.
	(Qu 20) 59%	majority of complaints (67%) refer to headaches, drzzmess, visual difficulties and mood alterations. Gastrointestinal problems (24%) and	
		affergie symptoms such as hives, rashes and swelling of tissues (15%) have also been reported. On occasion servires have been indeed with aspartance exposure. In most instances these difficulties were noted only with anisonits of aspartation exceeding normal use.	- What kind of seizures?

(PreTest

PostTest)

Double blind challenges have been carried out with aspariame. At Duke University, in one of the best designed such studies⁶, the effects of a single large dose of aspariame in people who had claimed to be sensitive to the substance was investigated. The results showed no difference in headache frequency, blood pressure, or blood histamme concentrations (a measure of allergenic potential) between the experimental and control groups, in another study at the investigy of llinois investigated. The results showed no difference in headache frequency, blood pressure, or blood histamme concentrations (a measure of allergenic potential) between the experimental and control groups, in another study at the investigy of llinois investigated. The results of the experimental and control groups in the placetor' group actually had more reactions than those in the aspariame group On this other hand, surveys by physicients in the adacetor' evolution precipitates headaches about 8% of the time. This kind of control aspariame Reported anecdotal experiences are not continued by calculity controlled accentitie studies. This of course does not mean that the problem sare not real, but does imply that in many cases the symptonic may not be caused by aspariame. People get beakaches, upset somaches, aches and paris of all kinds on a regular basis for no easily determined reason. If they recall having consumed aspariame when one of the adverse publicity that aspariame has received. At least one study has, however, construct allergies synptoms such as hives and swelling mis institue individuals. It is unclear how the allergy comes about, since noire of the components of adjustantine are believed to be capable of producing allergies reactions. It has been suggested that dide topperazine⁶, a compound which forms when aspariame decomposes may be reaptionable.

(Qu 21) 82%

A number of theoretical possibilities have been advanced to account for separtame associated problems. The three breakdown products of aspartame, are all toxic in high dores. Phenylalannet' is an essential annovaled which must be included in the dief for normal growth and maniferance, but sustained high blood levels can lead to brain damage. This is of major concern to the one out of roughly twenty thousand children which must be an ended condition called. The brain subdiment who are been with an inductive decondition called plantile damage of PST. These shifteen annot metabolize plantile damage of PST. These shifteen annot metabolize plantile damage of PST. These shifteen annot metabolize plantile damage which the builds up to damage to its levels in their boards the conduct interclote the conducts a solver containent of plantile in the robot the conducts a solver containent of plantile conducts are on the conduct of the tried solvers of his line and the conduct interclote the conduction of the tried solvers of his line and the conduct of the conduction of the output is consume the fluid flood operative solution and the conduction of the tried solver of the conduction and the conduction of the output is consumed being the conduction of the conduction of the output is consumed by the conduction of the conduction of the output is consumed by the conduction of the conduction of the output is consumed by the conduction of the conduction of the output is output in the fluid flood of output is a conduction of the output is output in the output is output in the output is a conduction of the output is output in the flood of output is a conduction of the output is output in the output is output in the output is a conduction of the output is output in the outpu

(Learner Commente)

200

5

6

- Define "double blind" (twice)

- Awkward wording (twice)

- *Define

- When something is ingested, does it break down? Is

decompose the same as break down?

- "Pronunciation?

-So what?

- "Is this a breakdown product of aspartame?

Pronunciation?

-This makes me feel dumb

7

8

-This sounds as if the PKU children are taking 100 sweeteners a day.

-So is 100 tablets a day dangerous or not?

-What unloward effects?

-Carbohydrates should be in a new paragraph This is jumping

(Learner Commenta)

around too much.

and back to phenylalanine.

- How can blood levels be raised?

- How is it eliminated?

- Too elaborately expressed.

- What's a scientific accusation?

"-This is the summary for the whole article

-'What does this phrase mean?

(Learner Comments)

- Table is very contusing

- Takes time to understand

- I don't know what each symbol means, but I have a general

idea

- I panicked when I saw the table.
- Why don't they include aspartame?
- They should give a table for comparing all the sweeteners

(twice).

In the general population, phenylalanine levels in the blood after aspartame ingestion are in the same range as after eating any protein containing food. Even at abusive amounts, equivalent to a child swallowing 100 sweetner tablets, levels do not rise above those which are considered to be safe in children afflicted with PKU Dr. Richard Wurtnan, a noted NIT researcher, has suggested that some of the unitoward effects of aspartame may be caused by a solden increase in brain phenylalanine levels, especially when the sweetner is consumed along with tools high in carbolydrates. Carbolydrates tingger insulin release into the bloodstream, which in turn makes it easier for phenylalanine to be absorbed by the brain. According to Wurtman, such a sudden increase in brain phenylalanine levels can cause depression, sleep problems, headaches and even seizures. These ideas have not been continued in human studies, and Wurtman, who as subjet mericas logities, and subjet molerately himself, feels the problem is only significant when consumption of aspartame is unnecessanly high.

The effects of asparace acid, another aspartame breakdown product, have also been rigorously examined. Administration of extremely large ansonits to non-human primates produced no damage even though blood levels were greatly elevated. In humans, even high dones are quickly climinated Nissi significantly, aspartic acid levels in the blood are not increased after eating aspartame containing loads or when drinking sweetened heverages even at the rate of three drinks in four human.

Perhaps the most unscientific accusations leveled at aspartame have involved its incluance content. It is a fact that in large doises methanol can lead to blindness and even to death. Alarmists have therefore referred to the methanol which is refersed from aspartame as an unsafe substance. In most be remembered, however, that there are no safe substances, only safe doeses 44. The amount of methanol which can be referased from aspartame is inconsequential in context of the overall deet. Methanol occurs naturally in foods. In fact the "instantal" methanol content of fruit piece is about 2.5 times that released in the body from a deet drink using aspartame as the sweetening agent. Even at the 99th centre level of 44 mg per kg bodyweight per day*, blood levels of methanol are not detectable.

In an a

(Qu 8) 71% (Qu 8) 76% (Qu 6) 71% (Qu 6) 65% (Qu 23) 88% (Qu 22) 76%

PostTest)

CYCLAMATE. SACCHARIN t S FA -FA + TTS -TTS + FA -CANADA FA -TTS + TTS * - = RANNED + = AVAILABLE FA = FOOD ADDITIVE: *- supermarkets, + drug stores restaurants

AVAILABILITY OF CYCLAMATE AND SACCHARIN

TTS = TABLE TOP SWEETIENER

(PreTest

(PreTest

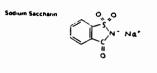
10

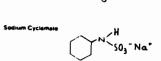
(ProTest

PostTest)

(Learner . Commenta)

Aspensme HOOC CH₂CHC -N-CH--CH₂--HOOC CH₂CHC -N-CH--CH₂--NH₂ Co₂CH₃





- It seems as if there's a title missing.

- This is OK for a science student; it means nothing to me.

- I don't know what they mean, but they look nice.

- I don't understand any of this.



(PreTest

PostTest)

.

(Learner Commente)

GENERAL COMMENTS OF LEARNERS

- It should start each section with a purpose or a finding, back it

up, repeat it.

- The science words are difficult for people with no background

in science.

- I can't pronounce the words.

-There's enough difficult material here without making the

language complicated too.

-There's no structure to this. I couldn't give an overview.

Structure it in comparative form. Start talking about sweeteners

in general, then each one.

- I wouldn't want to be tested on this.
- I'll think twice about buying a Diet Coke. It forces you to think.
- Important information needs recapping. It needs a conclusion.
- They never really explain what a sweetener is.
- The writers are assuming we know all this already.
- It needs more examples, more background information.

Subject Matter Expert (Diet Cancer Relationship) ^[13 (M1)]

SUBJECT'S QUALIFICATIONS

Degrees:

B Sc (Nutrition)

Professional Affiliations:

Corporation Professionel des Dietitists du Quebec Canadian Dietetics Association Quebec Dietetics Association.

Years of experience:

13 years.

Stated area of expertise:

General dietetics. Experience with the public as an audience.

Previous experience of reviewing teaching materials?

No.

Familiarity with content of materials? (1...2...3...4)

Familiarity with the intended audience of materials? (1...2...3...4)

DATA COLLECTION

Subject was audiotaped while reading through materials. The comments have been transcribed and summarized in the left hand column.

SUBJECT'S QUALIFICATIONS

Degrees:

B Sc (H Ec); Internship for professional certification.

Professional Affiliations:

Corporation Professionel des Dietitists du Quebec Canadian Dietetics Association Organization for Nutritional Education.

Years of experience:

17 years.

Stated area of expertise:

Dietician with a concentration on education in a non-classroom setting.

Previous experience of reviewing teaching materials?

No. Much experience with reviewing legislation.

Familiarity with content of materials? (1...2...3...4)

Familiarity with the intended audience of materials? (1...2...3...4)

DATA COLLECTION

Subject was audiotaped while reading through materials. The comments have been transcribed and summarized in the left hand column.

Target Population Expert (Diet-Cancer Relationship)^[14 (M1)]

SUBJECT'S QUALIFICATIONS

Degrees:

PhD in Organic Chemistry

Professional Affiliations:

None given

Years of experience:

16 years University and CEGEP teaching.

Stated area of expertise:

Pedagogy, Science

Previous experience of reviewing teaching materials?

Yes

Familiarity with content of materials? (1...2...3...4)

Familiarity with the intended audience of materials? (1...2...3...4)

DATA COLLECTION

Subject was audiotaped while reading through materials. The comments have been transcribed and summarized in the right hand column.

Target Population Expert (Artificial Sweeteners) [14 (M2)]

SUBJECT'S QUALIFICATIONS

Degrees:

B Sc; B Ed; BA; M Ed

Professional Affiliations:

National Science Teachers' Association; National Council of Teachers of Mathematics; Quebec Association of Mathematics Teachers

Years of experience:

25 years.

Stated area of expertise:

Teaching and therefore sensitivity to the way people learn.

Previous experience of reviewing teaching materials?

Yes

Familiarity with content of materials? (1...2...3...4)

Familiarity with the intended audience of materials? (1...2...3...4)

DATA COLLECTION

Subject was audiotaped while reading through materials. The comments have been transcribed and summarized in the right hand column.

Data Collection from Learners - (Diet Cancer Relationship)

[15 (M1)]

SUBJECTS

Seventeen students typical of the intended audience for the materials.

DATA COLLECTION

1. Test data

All subjects wrote a pretest, studied the materials, then wrote a posttest. Each test was based on key concepts identified by experts.

Questions	Pretest	Posttest
1 - 9	9 true/false	9 true/false (matched to pretest)
10 - 23	none (to avoid prelearning)	14 multiple choice
	(% of learners with correct answers - Column 1)	(% of learners with correct answers - Column 2)

2. Oral Comments

Three groups of learners were also audiotaped while they were studying the materials:

1 single learner	= 1
1 pair of learners	= 2
1 group of 5 learners	= 5
Total	= 8

Their comments were transcribed and summarized in the last column.

Where more than one learner made the same comment, this has been noted in brackets.

Data Collection from Learners - Artificial Sweeteners

[15 (M2)]

SUBJECTS

Seventeen students typical of the intended audience for the materials.

DATA COLLECTION

1. Test data

All subjects wrote a pretest, studied the materials, then wrote a posttest. Each test was based on key concepts identified by experts.

Questions	Pretest	Posttest
1 - 8	8 true/false	8 true/false (matched to pretest)
9 - 23	none (to avoid prelearning)	15 multiple choice
	(% of learners with correct answers - Column 1)	(% of learners with correct answers - Column 2)

2. Oral Comments

Three groups of learners were also audiotaped while they were studying the materials:

1 single learner	= 1
1 pair of learners	= 2
1 group of 5 learners	= 5
Total	=8

Their comments were transcribed and summarized in the last column.

Where more than one learner made the same comment, this has been noted in brackets.

Please read each question carefully and circle the <u>best</u> answer. After you answer each question, please rate how confident you feel about your answer on the scale provided.

TRUE FALSE QUESTIONS

CONFIDENCE RATING



[16 (M1)]

- 1. T F Current National Research Council guidelines recommend a fat intake of about 15% of total calories.

0)			ş O
5.	т	F	Sunlight and toxic waste are environmental factors which may cause cancer.	رون من
6.	Т	F	Carcinogens in the diet may be reduced by avoiding burnt food.	14
7.	т	F	Regular consumption of vegetables can reduce the risk of cancer by 50 percent.	14
8.	т	F	Many foods contain mutagens and natural anti-carcinogens.	14
9.	т	F	Although per capita correlation of dietary fats with cancer is strong, no conclusive correlation has been shown between individual fat consumption and cancer.	14
10.	т	F	Vitamin C is the major "free-radical trapping" anti- oxidant in human blood.	14
11.	т	F	Colon cancer has been associated with high cholesterol diets.	14
12.	Т	F	There is no evidence the either Vitamin C or Vitamin E can prevent cancer.	14
13.	т	F	According to a 1975 study: A negligible decrease in lung cancer rates among Norwegian men was obtained by increasing Vitamin A consumption.	14

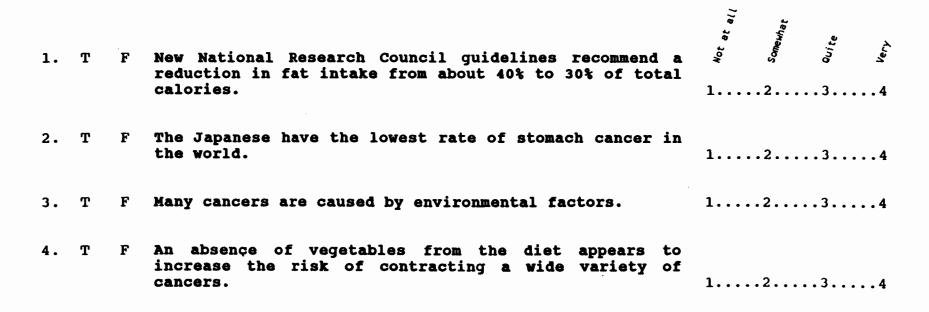


INSTRUCTIONS

Please read each question carefully and circle the <u>best</u> answer. After you answer each question, please rate how confident you feel about your answer on the scale provided.

TRUE FALSE QUESTIONS

CONFIDENCE RATING



0				
5.	т	F	Consumption of naturally occurring carcinogens is an environmental factor which may cause cancer.	۲
6.	т	F	Consumption of burnt food does <u>not</u> add carcinogens to the diet.	14
7.	T	F	The risk of cancer can be reduced by one third if vegetables are eaten regularly.	14
8.	т	F	Anti-carcinogens and mutagens rarely occur in the same food.	14
9.	Т	F	Individual fat consumption is highly correlated with cancer.	14
10.	Т	F	Vitamin E is the major "free-radical trapping" anti- oxidant in human blood.	14
11.	Т	F	No association exists between colon cancer and a diet high in cholesterol.	14
12.	T	F	Evidence indicates that Vitamin C and Vitamin E prevent cancer.	14
13.	Т	F	A 1975 study of Vitamin A consumption among Norwegian men showed the following: Increasing Vitamin A consumption decreased the rate of lung cancer by more than 50%.	14

MULTIPLE-CHOICE QUESTIONS

Not at all Somewhar ²uite Very 14. Controlled experiments repeated many times: 1....4 Prove theories a) **b**) Are the hallmark of the scientific method Are as valid as the intuition of famous scientists **c**) Must be conducted in laboratory settings. d) 1....4 15. Experiments in nutrition are difficult to carry out because: It is too difficult to find human subjects. a) It is difficult to provide "proof" based on results b) Governments will not fund the research. C) They require too much time and effort to be worthwhile. d) 16. Consumption of which family of vegetables is highly recommended in the National Research Council's nutritional 1.....4 quidelines? a) Cabbage b) Lequme Tuber C) d) Herb 17. Which of the following was not recommended by the National 1.....4 **Research** Council? A reduction of fat intake a) A reduction of the consumption of cured, pickled and b) smoked foods Vitamin supplementation C) An increased consumption of whole grain cereals, fruits d) and vegetables



18. Nutritional guidelines for the population as a whole:

- a) Neglect regional differences in the food availability making their application impossible.
- b) May be improperly applied leading many scientist to object to their recommendation.
- c) Do not consider foods from differing cultural origins.
- d) Ignore individual differences in height, weight and bone size.



- a) Smoking, drinking and not eating vegetables
- b) Smoking and eating meat

in the highest risk of cancer?

c) Smoking, drinking, eating meat and eating vegetables

19. Which of the following combinations of factors would result

d) Smoking, drinking and eating meat

20. Carotene, Vitamin C and Selenium are:

- a) Are all found in a glass of orange-juice
- b) Increase the speed of feces elimination
- c) All examples of anti-carcinogens
- d) Are found in DNA
- 21. Professor Bruce Ames of the University of California at Berkeley has devised a scale of relative risk (the H.E.R.P.) index. It is associated with:
 - a) The risk of contracting various forms of the Herpes Simplex virus.
 - b) The likelihood of encountering naturally occurring cancer-causing agents found in the environment on a given day.
 - c) Getting a variety of diseases and stands for High Environmental Risk Position.
 - d) Various substances in food and other parts of the environment in terms of the likelihood of getting cancer.

1.....4

1.....4

a)

b)

C) d)

a)

b)

C) d)

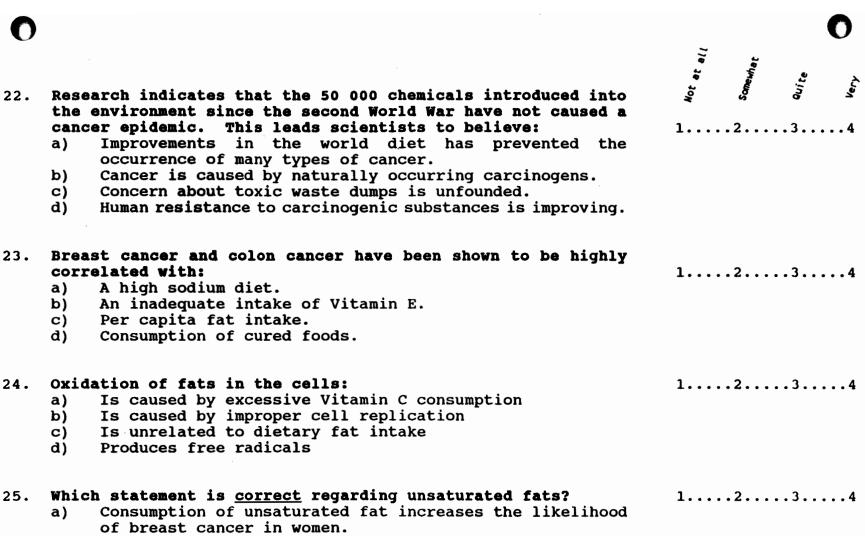
a)

b)

C)

d)

a)



- Unsaturated fats may pose an increased risk of cancer as b) they are easily oxidized.
- C) Unsaturated fats are environmental anti-carcinogens.
- d) Consumption of unsaturated fats is the cause many forms of cancer.

Not at all Somewhar Quite 26. Which of the following substances will not prevent the 1....4 production of free radicals? Vitamin A a) Beta carotene b) Vitamin E C) d) Selenium 27. Cancer of the stomach and esophogas are associated with: A high cholesterol diet. a) A high fat diet. b) Consumption of cured and smoked foods. C) Consumption of red meat. d) 28. Selenium is a mineral required in trace amounts. What is its 1.....4 function? It prevents cancer of the esophogas. a) It is required in the digestive process. b) It assists in metabolizing beta carotene. C) It protects cells from damage by oxidation. **d**) 1....4 Which of the following is true about Vitamin A: 29. The precursor to Vitamin A, beta-carotene, is thought to a) be the important factor in reducing some forms of cancer. Intake of Vitamin A supplements was highly recommended b) by the National Research council to prevent many forms of cancer. Vitamin A is not ingested by the human body. C) Vitamin A is not toxic, even at high doses. d) The probability of contracting cancer of the small intestine 1....4 30. and colon may be reduced by: An increase in Vitamin C consumption a) Eliminating BHT from the diet. b) Increasing consumption of dietary fiber. C) Increasing consumption of dairy products. d)

INSTRUCTIONS

Please read each question carefully andcircle the best answer. After you answer each question, please rate how confident you feel about your answer by circling the appropriate number on the scale, where 1 means "not confident", 4 means "very confident"

				CONFIDENCE RATING
				Not at all somewhat outre Jery
1.	Saccharin was discovered by accident.	т	F	14
2.	"Safe under all conditions" and "Generally Recognised as Safe" mean the same thing.	т	F	1
3.	Canadian studies in 1977 showed that saccharin caused bladder cancer.	т	F	1
4.	Cyclamates are similar in sweetness to sugar.	т	F	14
5.	Cyclamates are not dangerous but cyclohexamine is.	т	F	14
6.	The amount of methanol produced when aspartame decomposes is so small that it is not dangerous	т	F	14
7.	Complaints about adverse reactions to aspartame are often heard in North America	т	F	14
8.	Aspartic acid is a toxic breakdown product of aspartame	т	F	14

16 (M2)

CONFIDENCE RATING

INSTRUCTIONS

Please read each question carefully andcircle the best answer. After you answer each question, please rate how confident you feel about your answer by circling the appropriate number on the scale, where 1 means "not confident", 4 means "very confident"

				NOT at all Somewhat Outre Very
1.	Saccharin was discovered after a long and deliberate search	т	F	14
2.	The Delaney Act applies to substances classified as "Generally Recognized as Safe"	т	F	14
3.	In 1977, the Canadian Government banned the use of saccharin as an additive	т	F	14
4.	Cyclamates are much sweeter than sugar	т	F	14
5.	Cyclamates break down into cyclohexamine in the body	т	F	14
6.	Aspartame is dangerous because it breaks down to give methanol	т	F	14
7.	Complaints about adverse reactions to aspartame are associated with excessive use of aspartame	т	F	14
8.	Aspartic acid breaks down to give aspartame	т	F	14

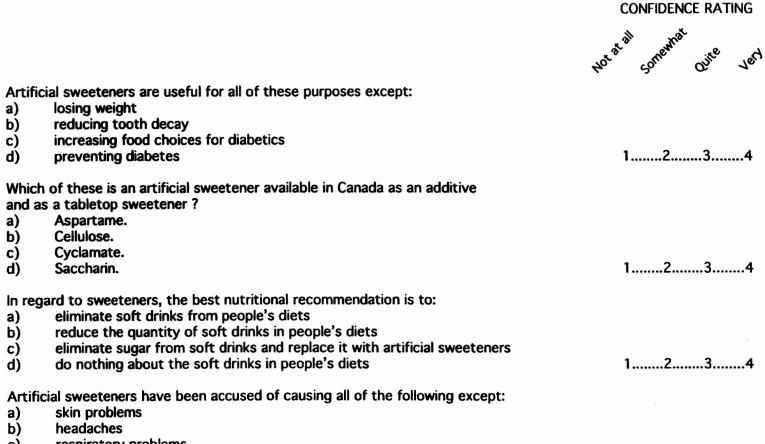
9.

10.

11.

12.

Choose the best answer in each case



- c) respiratory problems
- d) diabetes

1.....4



13.	 Which of these statements about artificial sweeteners is true? a) Used in modest amounts, they represent no danger. b) Used in abusive amounts, the risk is increased, but still low. c) The risk-benefit ratio associated with them means that they are safe. 	
	 d) They allow diabetics to eat any food they choose. 	1
14.	Which of these statements about saccharin is true ?	
	a) It has been used in the United States for over 100 years.	
	b) It is twice as sweet as sugar.	
	 It is metabolized by the body and has caloric value. 	
	d) It has been proved to cause bladder cancer in humans.	14
15.	Which of these statements about saccharin is true ?	
	a) It is 300 times sweeter than sugar and adds "body" to beverages.	
	b) It is not absorbed by the body and has good "mouthfeel".	
	c) It adds "body" to beverages because of its good "mouthfeel".	
	d) It is not absorbed by the body and is 300 times sweeter than sugar.	14
16.	The best evidence available about saccharin is that it:	
	a) causes cancer at low dose rates	
	b) causes cancer in cells already exposed to carcinogens	
	c) does not cause cancer	
	d) causes cancer by decomposing into proteins	1
17.	Which of these statements about cyclamates is true ?	
	a) They have been proved to cause cancer in humans.	
	b) They are more stable than saccharin or aspartame.	
	c) They cause the testes of rats to atrophy.	
	d) They are excreted by the body without being broken down.	1



18.	 Which of the following statements regarding aspartame is true ? a) It is about as sweet as an equal weight of sugar but contains fewer calories. b) It is approved for use in Canada as a food additive. c) It is non-caloric and cannot be metabolized. d) It has a sugar-like structure. 	14
19.	 Which of the following statements regarding aspartame is true ? a) Aspartame breaks down in the body to give phenylalanine, aspartic acid and methanol. b) Aspartame is a useful substitute for sugar in cooking and baking. 	14
	 c) When aspartame is given in high doses the aspartic acid level in the blood goes up sharpl d) Aspartame cannot be metabolized by the body. 	y. 1234
20.	 Which of the following statements is true regarding research on aspartame ? a) Few studies have been done, with few reports of adverse effects. b) Carefully controlled studies have confirmed anecdotal reports of problems. c) Double blind studies have shown no adverse reactions. d) No studies exist to show an allergic reaction to aspartame. 	14
21.	Aspartame should not be consumed in excess by children suffering from: a) diabetes b) high blood pressure c) phenylketonuria d) diarrhea	14
22.	 In order of increasing sweetness, the correct order is: a) sugar, cyclamates, aspartame, saccharin b) saccharin, cyclamates, sugar, aspartame c) aspartame, sugar, cyclamates, saccharin d) cyclamates, aspartame, saccharin, sugar 	14

.



- 23.
- "There are no safe substances, only safe doses" means: a) All substances are safe, only large doses are dangerous. b) Any substance is dangerous if it is abused.

 - c)
 - Allowed sweeteners are safe at any dosage. All sweeteners should be used in small doses only. d)

1......4

THANK YOU VERY MUCH

.

Possible prompting words during Think-aloud: [17]

- 1. Keep thinking aloud, please.
- 2. Keep talking, please.
- 3. Tell me what you're thinking, please.
- 4. Tell me more about it.
- 5. Is there something you want to say about that?
- 6. Tell me more
- 7. What are you thinking, please ?
- 8. Please remember to keep talking

Last resort:

9. Is something wrong?

Debriefing Script [19 (T1) w]

(Tape recorder should be running)

Thank you for taking part in this study. As you were working through the materials, I took some notes. I wasn't permitted to ask or answer any questions then, but I'd like check some things that were unclear to me, then ask a few more general questions.

(Ask questions generated by tracking sheet re: unclear comments, data sources for unlinked revisions, any other issues)

Thank you for clarifying these points.

Now some more general questions:

1. How recently have you been involved in the revision of materials?

2. What were the materials?

3. How close was today's task to your usual procedure when you revise instructional materials? What were the similarities and differences ?

4. When you revise instructional materials, do you usually have access to external data? If so, what data?

Do you focus on a particular data source ? If so, which one(s) (e.g. -- if needed -- learner, expert or your own experience).

5. In today's task, did you focus on one data source more than another in terms of the time you spent with it?

In terms of the amount of attention you paid to it? Why?

How did you establish priorities?

6. In this task, what did you focus on while you were reviewing the materials (If needed): How did you establish priorities?

Did you use a particular area of your expertise?

What criteria did you keep at the back of your mind?

7. Did you experience any problems with the task you were given (e.g. the time involved, having to verbalize all your thoughts, etc)?

8. If you go back to the office and have to describe this task to a colleague, how would you describe the task you've just completed ?

9. Would you like to give some overall impression or opinion of the materials?

10. How would you rate your familiarity with the content matter of the module? (1...2...3...4)

11. How would you rate your familiarity with the target audience for the module? (1...2...3...4)

12. Do you have any other comments?

13. Are there any questions that I could answer for you?

Debriefing Script [19 (T1) n]

(Tape recorder should be running)

Thank you for taking part in this study. As you were working through the materials, I took some notes. I wasn't permitted to ask or answer any questions then, but I'd like check some things which were unclear to me, then ask a few more general questions.

(Ask questions generated by tracking sheet)

Thank you for clarifying these points.

Now some more general questions:

1. How recently have you been involved in the revision of materials?

2. What were the materials?

3. How close was today's task to your usual procedure when you revise instructional materials ? (1...2...3...4)

What were the similarities and differences ?

4. When you revise instructional materials, do you usually have access to external data? If so, what data?

Do you usually focus on a particular data source ? If so, which one(s)

(e.g. -- if needed -- learner, expert or your own experience).

5. In today's task, How did you go about revising the materials ?

(If needed: How did you establish priorities?

Did you use a particular area of your expertise?

What criteria did you keep at the back of your mind?

6. Did you experience any problems with the task you were given (e.g. the time involved, having to verbalize all your thoughts, etc)?

7. If you go back to the office and have to describe this task to a colleague, how would you describe the task you've just completed ?

8. Would you like to give some overall impression or opinion of the materials?

9. How would you rate your familiarity with the content matter of the module? (1...2...3...4)

10. How would you rate your familiarity with the target audience for the module? (1...2...3...4)

11. Do you have any other comments ?

12. Are there any questions that I could answer for you?

Debriefing Script [19 (T2) w]

(Tape recorder should be running)

Thank you for taking part in this study. As you were working through the materials, I took some notes. I wasn't permitted to ask or answer any questions then, but I'd like clarify some things that were unclear to me, then ask a few more general questions, then give you the chance to ask me questions.

(Ask questions generated by tracking sheet re: unclear comments, data sources for unlinked revisions, any other issues)

Thank you for clarifying these ponts.

Now some more general questions:

1. You said last time that you had recently been involved in revising

3. How close was today's task to your usual procedure when you revise instructional materials?

What were the similarities and differences ?

4. You also said last time that you have access to external data sources and that you usually focus on the following data:

Which one do you usually pay most attention to in terms of the time you spend on it?

In terms of the amount of importance you give to it?

5. In this task, did you focus on one data source more than another in terms of the time you spent on it?

In terms of the amount of attention you paid to it?

Why?

How did you establish priorities?

6. Did you use a particular area of your expertise?

What criteria did you keep at the back of your mind?

7. Did you experience any problems with the task you were given (e.g. the time involved, having to verbalize all your thoughts, etc)?

8. If you go back to the office and have to describe this task to a colleague, how would you describe the task you've just completed ?

9. Would you like to give some overall impression or opinion of the materials?

10. How would you rate your familiarity with the content matter of the module? (1...2...3...4)

11. How would you rate your familiarity with the target audience for the module? (1...2...3...4)

12. Do you have any other comments ?

13. Are there any questions that I could answer for you?

Now I'm going to ask you a few questions on your background.

1. Would you give me a full list of your professional qualifications (e.g., degrees, professional affiliations etc?

- 2. Where did you study ID?
- 3. At university, how many courses did you take in instructional design?
- 4. What were they?
- 5. Did you take any courses specific to formative evaluation or to revision?
- 6. What textbooks did you use in ID courses?
- 7. What text material, if any, do you still refer to?
- 8. What professional journals do you read?
- 9. What conferences have you attended lately?
- 10. How many years of experience do you have in your field ?

Debriefing Script [19 (T2) n]

(Tape recorder should be running)

Thank you for taking part in this study. As you were working through the materials, I took some notes. I wasn't permitted to ask or answer any questions then, but I'd like check some things which were unclear to me, then ask a few more general questions, then give you a chance to ask me questions.

(Ask questions generated by tracking sheet)

Thank you for clarifying these points.

Now some more general questions:

1. You said last time that you had recently been involved in revising:

How close was today's task to your usual procedure when you revise instructional materials?

2. What were the similarities and differences ?

3. You also said that you usually have access to external data and that you usually focus on the following data:

In today's task, what did you focus on while you were reviewing the materials? (If needed): How did you establish priorities?

4. Did you use a particular area of your expertise?

5. What criteria did you keep at the back of your mind?

6. Did you experience any problems with the task you were given (e.g. the time involved, having to verbalize all your thoughts, etc)?

7. If you go back to the office and have to describe this task to a colleague, how would you describe the task you've just completed ?

8. Would you like to give some overall impression or opinion of the materials?

9. How would you rate your familiarity with the content matter of the module? (1...2...3...4)

10. How would you rate your familiarity with the target audience for the module? (1...2...3...4)

11. Do you have any other comments ?

12. Are there any questions that I could answer for you?

Now I'm going to ask you a few questions on your background.

1. Would you give me a full list of your professional qualifications (e.g., degrees, professional affiliations etc?

2. Where did you study ID?

- 3. At university, how many courses did you take in instructional design?
- 4. What were they?

- 5. Did you take any courses specific to formative evaluation or to revision?
- 6. What textbooks did you use in ID courses?
- 7. What text material, if any, do you still refer to?
- 8. What professional journals do you read?
- 9. What conferences have you attended lately?
- 10. How many years of experience do you have in your field ?

That's all for today!

Thank you very much for the time and effort you've put into this task. I know that doing a Think-aloud is a demanding process and I hope it hasn't been too tiring for you.

I think your experience with this session will make the next one much less tiring.

Can we confirm the time and place of the next session ? (Do so)

When you come back on ______, there'll be another, similar, module to revise, under different conditions. At the end of the next session, you may have some more questions, and we'll have an honorarium for you.

Until then, please accept our thanks.

<u>Goodbyes</u> [20 (T2)]

That's all!

I feel we have a lot of very useful data here, and I appreciate the effort that you put into it. I know that doing a Think-aloud is a demanding process and I hope it hasn't been too tiring for you.

Once again, thank you for taking the time to help us with this research. Please accept this as a token of our thanks. I'll need one more signature from you to acknowledge receipt of the cheque.



Faculty of Education McGill University 3700 McTavish Street Montreal, PO, Coast Montreal PQ, Canada - H3A 1Y2

Faculte des sciences de Feducation — Facsimile/Telecopier Universite McGill — 514: 398:4679 3700 rue McTavish Montrez IPD, Canada - H34,110

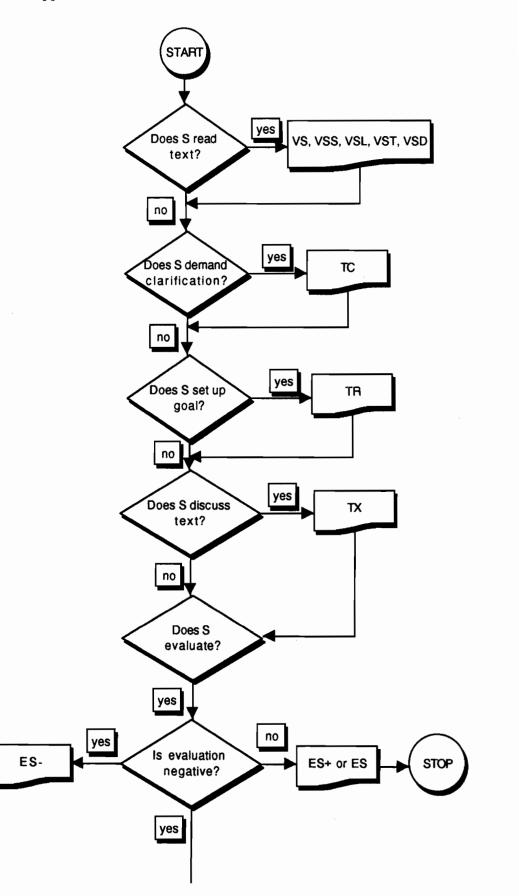
Receipt for honorarium

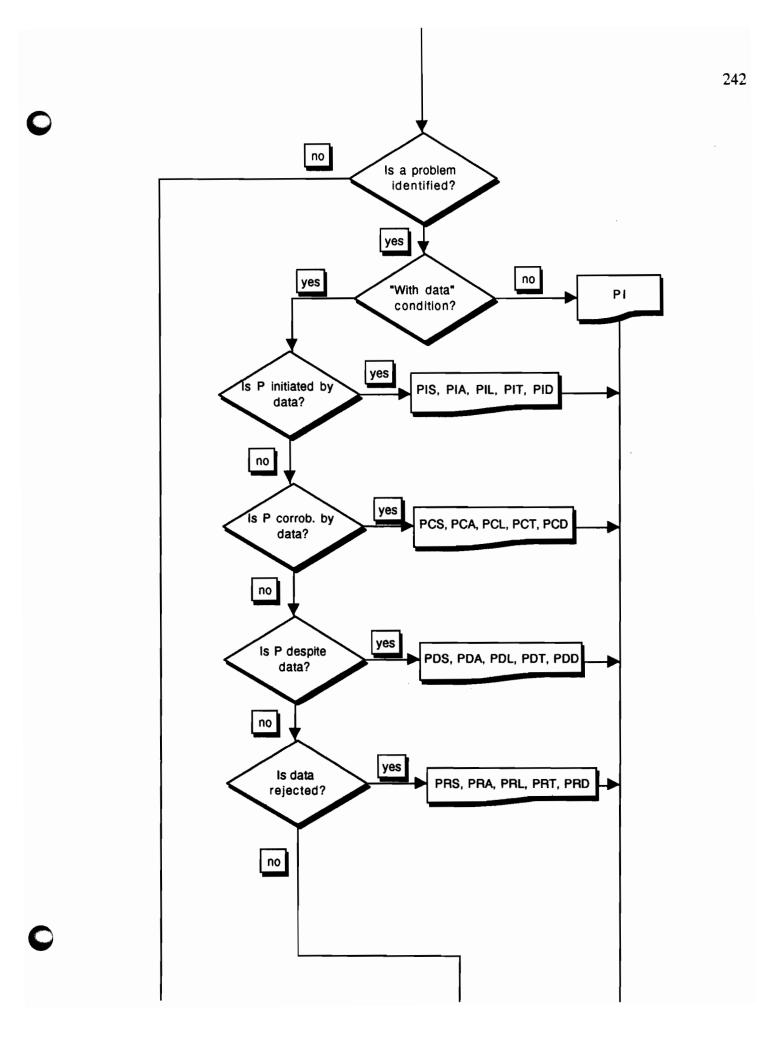
I acknowledge receipt of the sum of \$300 as honorarium for the time spent reviewing materials as part of the research project "Expert Review of Instructional Materials"

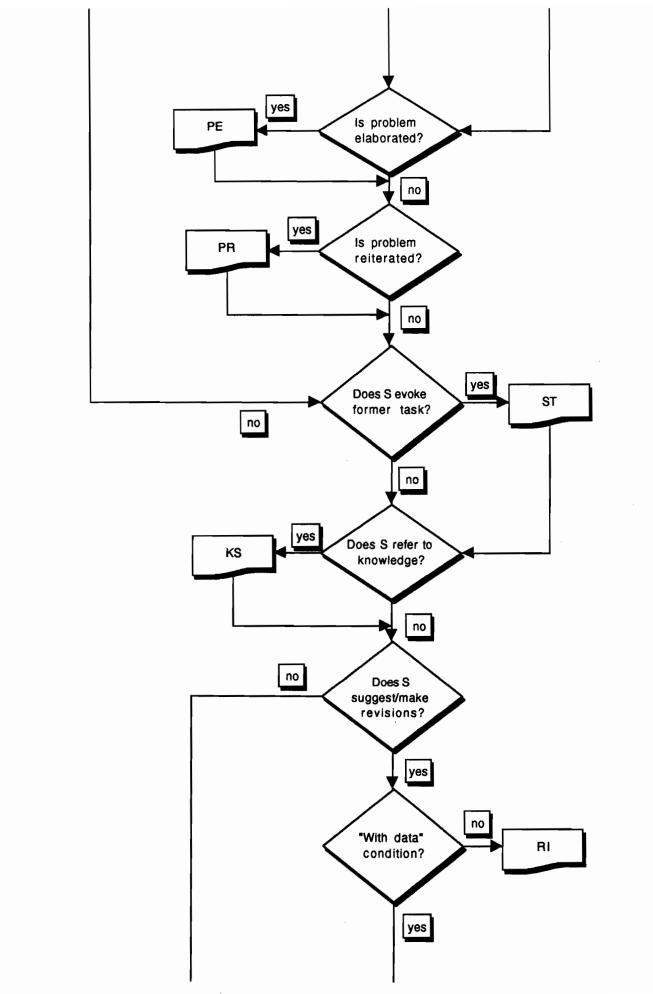
> Name: Signature:

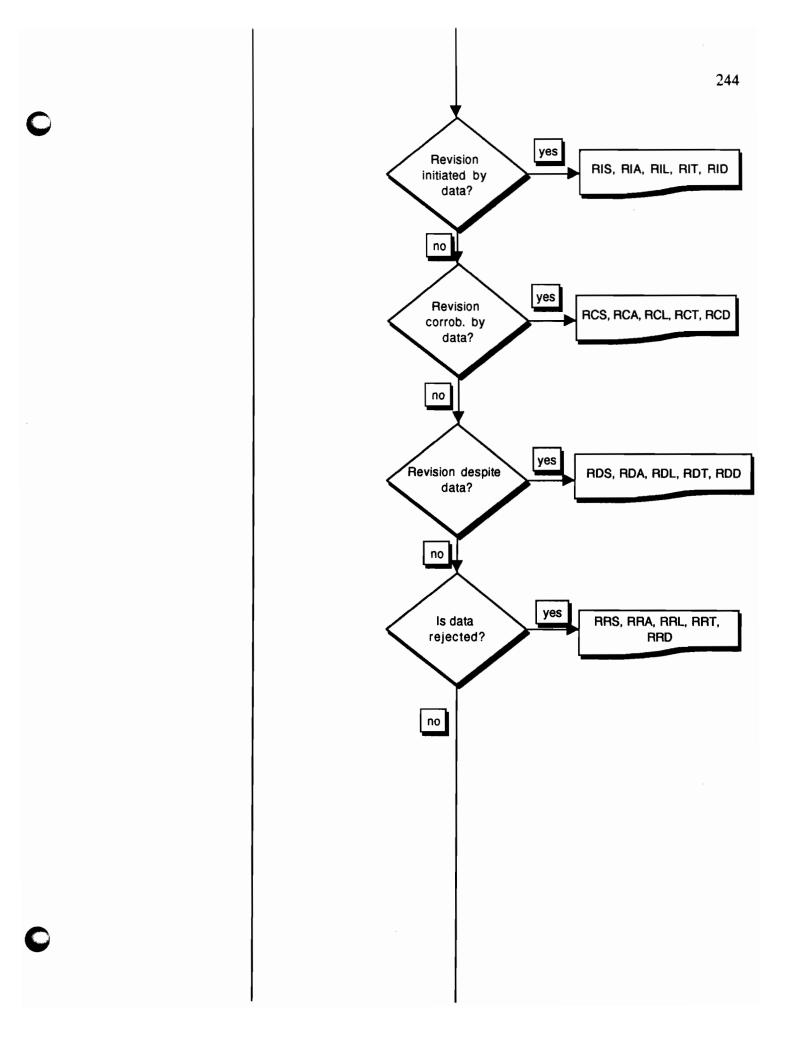
Date:

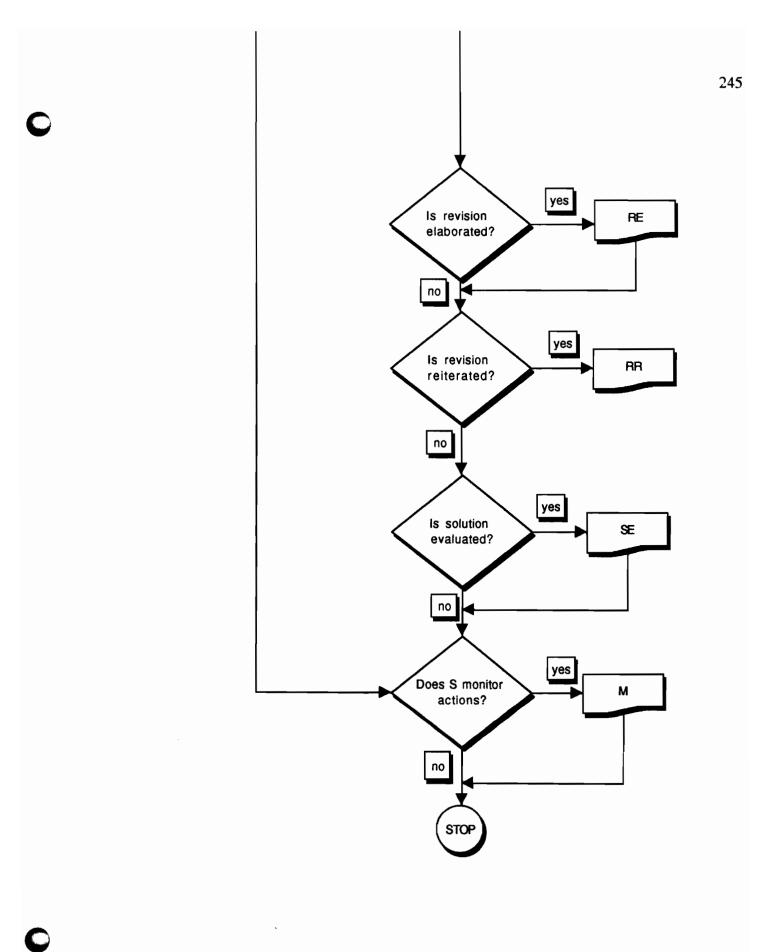
,



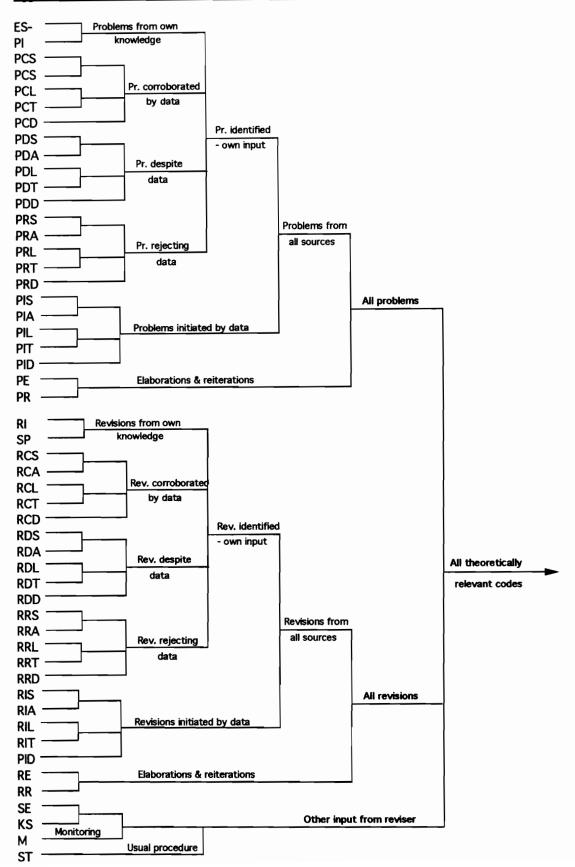






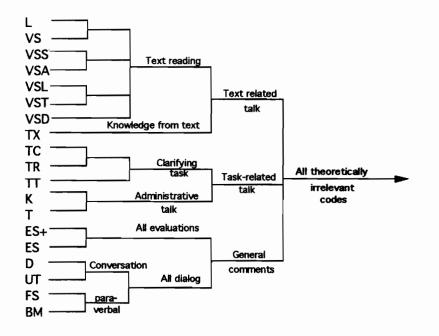


Appendix D: DENDROGRAM SUMMARIZING THEORETICALLY RELEVANT CODES



Appendix D (contd)

DENDROGRAM SUMMARIZING THEORETICALLY IRRELEVANT CODES



Appendix E. 72 categories of coded segments

ES-24324682081511152617263065PI4418611879873115486572455384532PIS001001010360800101PIA160100702130000000PIT0000000101000000PIT0000010010000000PIT00000001001000000PIT0000000100000000000000000000000000000000000000000000000000000000000000 <th>S17</th> <th>Г1 S</th> <th>1T2</th> <th>S2T1</th> <th>S2T2</th> <th>S3T1</th> <th>S3T2</th> <th>S4T1</th> <th>S4T2</th> <th>S5T1</th> <th>S5T2</th> <th>S6T1</th> <th>S6T2</th> <th>S7T1</th> <th>S7T2</th> <th>248 58T</th> <th>1 S8T2</th>	S 17	Г1 S	1T2	S2T1	S2T2	S3T1	S3T2	S4T 1	S4T2	S5T1	S5T2	S6T1	S6T2	S7T 1	S7T2	248 58T	1 S8T2
PI 44 18 61 18 79 87 31 15 48 65 7 24 55 38 45 32 PIS 0 0 1 0 1 0 3 6 0 88 0 0 1 0 1 PIA 16 0 1 0 0 2 13 0 23 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 <td></td>																	
PIA 16 0 1 0 0 1 4 0 3 0 0 5 0 4 PIL 13 0 0 0 0 0 0 0 1 0 2 13 0 23 0 0 10 10 PIL 1 0 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		44								48							
PIL 13 0 0 0 0 2 13 0 23 0 0 19 0 10 PIT 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	PIS	0	0	1	0	0	1	0	3	6	0	8	0	0	1	0	1
PIT 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	PIA	16	0	1	0	0	9	0	11	4	0	3	0	0	5	0	4
PID 1 0 0 4 0 4 0 0 0 0 0 PCS 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	PIL	13	0	0	0	0	7	0	2	13	0	23	0	0	19	0	10
PCS 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	PIT	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0
PCA30100201000000100PCL300000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000<	PID	1	0	0	0	0	1	0	0	4	0	4	0	0	4	0	4
PCL3000300600000000PCD0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000<	PCS	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	3
PCT0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000 <t< td=""><td>PCA</td><td>3</td><td>0</td><td>1</td><td>0</td><td>0</td><td>2</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>2</td><td>0</td><td>1</td></t<>	PCA	3	0	1	0	0	2	0	1	0	0	0	0	0	2	0	1
PCD0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000 <t< td=""><td>PCL</td><td>3</td><td>0</td><td>0</td><td>0</td><td>0</td><td>3</td><td>0</td><td>0</td><td>6</td><td>0</td><td>0</td><td>0</td><td>0</td><td>4</td><td>0</td><td>5</td></t<>	PCL	3	0	0	0	0	3	0	0	6	0	0	0	0	4	0	5
PDS000010100400100PDA200000000000100100PDL100000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000	PCT	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
PDA 2 0 0 0 1 0 0 1 0 0 1 0 0 PDL 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 </td <td>PCD</td> <td>0</td>	PCD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PDL 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		0	0	0	0	0		0	1	0	0	4	0	0	1	0	0
PDT 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		2				-		-				-		0	1		
PDD 2 0 0 0 0 0 0 2 0 0 2 0 0 2 0 0 PRS 4 0 0 0 0 6 0 2 3 0 1 0 0 2 0 2 PRA 3 0 0 0 0 9 0 2 4 0 2 0 0 0 0 3 PRL 4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 <td></td>																	
PRS 4 0 0 0 6 0 2 3 0 1 0 0 2 0 2 PRA 3 0 0 0 0 9 0 2 4 0 2 0 0 0 0 3 PRL 4 0 0 0 0 1 0 1 1 0 0 0 0 1 0 5 PRT 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0							-	-					-	0		0	0
PRA 3 0 0 0 9 0 2 4 0 2 0 0 0 0 3 PRL 4 0 0 0 0 1 1 1 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 </td <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>-</td> <td></td> <td></td> <td></td> <td>2</td> <td></td> <td></td> <td></td> <td></td> <td></td>								-				2					
PRL 4 0 0 0 0 1 0 1 1 0 0 0 1 0 5 PRT 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 </td <td></td> <td>-</td> <td></td> <td></td> <td></td> <td></td> <td></td>												-					
PRT 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0			-					-									
PRD 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 11 22 18 3 11 46 0 14 2 1 8 RIS 0 0 0 0 0 11 0 16 4 0 2 0 0 12 11 11 0 14 0 12 11 11 0 13 0 0 0 12 11 11 10 10 13 0 0 0 13 10 11 11 10 10 10 11 11 11 11 11 11 11 <td< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></td<>																	
RI 125 72 117 163 83 188 80 44 69 66 58 68 74 79 145 56 SP 6 0 1 0 0 1 22 18 3 11 46 0 14 2 1 8 RIS 0 0 0 0 2 0 1 11 0 7 0 0 3 0 9 RIA 10 0 0 0 11 0 16 4 0 2 0 0 4 0 12 RIL 23 0 1 0 0 3 3 0 0 0 13 0 RIT 2 0 0 0 3 0 3 3 0 0 0 14 0 12 RIT 2 0 0 0 3 0 3 0 0 0 1 145 0 12 <td></td> <td>-</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>-</td> <td>-</td> <td>-</td> <td></td> <td></td> <td></td>		-										-	-	-			
SP 6 0 1 0 1 22 18 3 11 46 0 14 2 1 8 RIS 0 0 0 0 2 0 1 11 0 7 0 0 3 0 9 RIA 10 0 0 0 11 0 16 4 0 2 0 0 4 0 12 RIL 23 0 1 0 0 9 0 15 10 0 3 0 0 27 7 RIT 2 0 0 0 3 0 3 0 0 0 3 0 3 RID 3 0 0 0 1 0 0 0 0 0 1 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3									-								
RIS 0 0 0 2 0 1 11 0 7 0 0 3 0 9 RIA 10 0 0 0 11 0 16 4 0 2 0 0 4 0 12 RIL 23 0 1 0 0 9 0 15 10 0 3 0 0 15 0 27 RIT 2 0 0 0 3 0 3 3 0 0 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 3 0 3 0<																	
RIA 10 0 0 11 0 16 4 0 2 0 0 4 0 12 RIL 23 0 1 0 0 9 0 15 10 0 3 0 0 15 0 27 RIT 2 0 0 0 0 3 0 3 0 0 0 3 0 3 RID 3 0 0 0 2 0 7 13 0 0 0 9 0 6 RCS 0 0 0 0 1 0 0 0 0 0 1 0 3 0 3 0 1 0 3 0 1 0 3 0 3 0 3 0 3 0 3 0 1 0 0 0 1 0 3 0 0 1 0 1 0 1 1 0 0 0 <t< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></t<>																	
RIL 23 0 1 0 0 9 0 15 10 0 3 0 0 15 0 27 RIT 2 0 0 0 0 3 0 3 3 0 0 0 0 3 0 3 3 0 0 0 3 0 3 3 0 0 0 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 3 0 0 0 3 0 3 0 3 0 3 0 3 0 3 10 3 3 0 0 0 1 0 3 0 0 1 0 3 3 3 <			•														
RIT 2 0 0 0 3 0 3 3 0 0 0 0 3 0 3 RID 3 0 0 0 2 0 7 13 0 0 0 9 0 6 RCS 0 0 0 0 1 0 0 0 0 0 1 0 3 RCA 0 0 0 0 5 0 0 1 0 0 0 0 0 1 0 3 RCA 0 0 0 0 5 0 0 1 0 0 0 0 1 RCL 2 0 0 0 5 0 0 4 0 0 0 1 0 1 0 1 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 0 </td <td></td>																	
RID 3 0 0 0 0 2 0 7 13 0 0 0 9 0 6 RCS 0 0 0 0 1 0 0 0 0 0 1 0 3 RCA 0 0 0 0 5 0 0 1 0 0 0 0 0 1 0 3 RCA 0 0 0 0 5 0 0 1 0 0 0 0 0 1 1 0 0 0 0 1 1 1 1 0 0 0 1 1 1 0 0 0 1 1 0 0 0 1 1 0 0 0 1 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0																	
RCS 0 0 0 1 0 0 0 0 0 0 1 0 3 RCA 0 0 0 0 5 0 0 1 0 0 0 0 0 1 0 3 RCA 0 0 0 0 5 0 0 1 0 0 0 0 0 1 RCL 2 0 0 0 5 0 0 4 0 0 0 10 0 1 RCL 2 0 0 0 5 0 0 4 0 0 0 1 1 0 0 0 1 0 1 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0																	
RCA 0 0 0 0 5 0 0 1 0 0 0 0 0 1 RCL 2 0 0 0 0 5 0 0 4 0 0 0 0 1 0 1 10 0 10 0 1 RCL 2 0 0 0 0 0 1 1 0 0 0 0 1 1 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 <																	
RCL 2 0 0 0 5 0 0 4 0 0 0 10 0 1 RCT 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0<	RCA		0														
RCT 0 0 0 0 0 1 1 0 0 0 1 0 0 RCD 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 </td <td>RCL</td> <td>2</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>5</td> <td>0</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>	RCL	2	0	0	0	0	5	0									
RCD 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	RCT	0	0	0	0	0	0										
RDS 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	RCD	0	0	0	0	0											
RDA 0 0 0 0 0 0 1 0 0 3 0 0 0 0 0			0														
	RDA	0	0	0	0	0											
	RDL	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1

Appendix E. 72 categories of coded segments (continued)

				-											
				•		•		•	0	•	~	•		249	•
RDT 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RDD 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RRS 1	0	0	0	0	1	0	1	2	0	1	0	0	2	0	2
RRA 1	0	0	0	0	13	0	4	4	0	7	0	0	0	0	5
RRL 5	0	0	0	0	6	0	1	2	0	0	0	0	0	0	2
RRT 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RRD 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RUS 0	0	0	0	0	1	0	0	0	0	0	0	0	3	0	0
SE 45	14	25 72	32	10	39 58	48	14 72	29	10	49 102	33	29	57	7	9
KS 65	21	73	53	38	58	61 42	72 52	153	41	103	28	16 20	60 06	29	28
M 42	21	19	12	3	11	42	53	11	6	6 22	10 45	30	96	4	11
ST 7	20	63 10	7	11	8	30 25	2	5	6 4	22	45 2	45	9	2 8	2 9
TC 9 TR 114	4 26	10 88	8 48	3 11	17 50	35 86	18 102	11 31	4	30 78	2 17	12 49	22 151	8 31	9 40
VS 91	20 39	00 123	40 134	81	233	80 48	102 39	78	45	78 66	93	49 122	116	99	40 84
VS 91 VSA44	39 0	125	0	0	233 70	40 0	90	78 18	43 0	101	93 0	0	45	99 0	0 4 48
VSL 29	0	23	0	0	33	0	90 46	18 66	0	64	0	0	45 119	0	103
VSE 25	0	25 6	0	0	25	0	32	24	0	66	0	0	33	0	49
VST 2	0	21	0	0	40	0	28	24 44	0	12	0	0	100	0	39
VST 2 VSD17	0	7	0	0	1	0	20	6	0	16	0	0	15	0	7
TX 179	54	, 137	109	<u>99</u>	127	70	39	137	40	218	278	145	209	101	, 67
PE 7	0	14	0	11	2	8	3	18	29	3	5	10	1	2	8
PR 8	0	13	1	8	9	10	8	26	11	3	4	23	5	3	3
RE 58	13	38	63	14	28	36	21		48	46	29	28	47	23	26
RR 51	16	44	72	4	51	37	19	35	18	22	33	30	98	30	16
TT 12	8	20	3	7	23	25	10	6	5	4	1	7	22	1	0
L 48	12	17	16	9	30	11	23	20	21	31	7	18	48	30	13
ES+ 16	21	9	4	13	6	27	8	15	5	10	15	25	17	5	4
ES 13	2	8	8	2	8	11	9	12	10	24	10	14	35	15	8
D 10	23	36	30	7	39	12	27	30	11	15	6	83	63	0	2
BM 13	12	8	14	10	12	3	12	19	2	40	6	45	42	2	8
FS 248	42	101	85	66	69	148	139	239	97	212	96	110	172	190	83
UT 22	7	124	84	4	33	15	11	73	21	66	0	5	71	0	11
K 61	59	81	57	37	313	98	111	130	98	96	32	62	216	33	42
T 1	3	0	1	0	2	0	.0	2	0	6	0	3	5	0	1
TOT152	27510	1322	1028	618	1736	1002	1111	1540	696	1617	859	1080	2120	812	933

Appendix F. 50 theoretically relevant coded segments

250

S1T1 S1T2 S2T1 S2T2 S3T1 S3T2 S4T1 S4T2 S5T1 S5T2 S6T1 S6T2 S7T1 S7T2 S8T1 S8T2

50 04	•			0	20	•	16	11	15	24	17	26	20		-
ES- 24 PI 44	3 18	24	6 18	8 79	20 87	8 31	15 15	11 48	15 65	26 7	17 24	26 55	30 38	6 45	5 32
PI 44 PIS 0	18	61 1	0	0	1	0	3	40 6	05	8	24 0	0	1	4 5 0	1
PIA 16	0	1	0	0	9	0	11	4	0	3	0	0	5	0	4
PIL 13	0	0	0	0	, 7	0	2	13	0	23	0	0	19	0	10
PIT 0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0
PID 1	0	0	0	0	1	0	0	4	0	4	0	0	4	0	4
PCS 0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	3
PCA 3	0	1	0	0	2	0	1	0	0	0	0	0	2	0	1
PCL 3	0	0	0	0	3	0	0	6	0	0	0	0	4	0	5
PCT 0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
PCD 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PDS 0	0	0	0	0	1	0	1	0	0	4	0	0	1	0	0
PDA 2	0	0	0	0	2	0	0	0	0	1	0	0	1	0	0
PDL 1	0	0	0	0	0	0	0	0	0	0	0	0	4	0	1
PDT 0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
PDD 2	0	0	0	0	0	0	0	0	0	2	0	0	2	0	0
PRS 4	0	0	0	0	6	0	2	3	0	1	0	0	2	0	2
PRA 3	0	0	0	0	9	0	2	4	0	2	0	0	0	0	3
PRL 4	0	0	0	0	1	0	1	1	0	0	0	0	1	0	5
PRT 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PRD 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RI 125	72	117	163	83	188	80	44	69	66	58	68	74	79	145	56
SP 6	0	1	0	0	1	22	18	3	11	46	0	14	2	1	8
RIS 0	0	0	0	0	2	0	1	11	0	7	0	0	3	0	9
RIA 10	0	0	0	0	11	0	16	4	0	2	0	0	4	0	12
RIL 23	0	1	0	0	9	0	15	10	0	3	0	0	15	0	27
RIT 2 RID 3	0 0	0	0	0	3 2	0	3	3	0 0	0	0	0	3	0	3
RCS 0	0	0 0	0 0	0 0	2	0 0	7 0	13 0	0	0 0	0 0	0 0	9 1	0 0	6
RCA 0	0	0	0	0	5	0	0	1	0	0	0	0	0	0	3 1
RCL 2	0	0	0	0	5	0	0	4	0	0	0	0	10	0	1
RCT 0	0	0	0	0	0	0	1	- 1	0	0	0	0	10	0	0
RCD 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RDS 0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RDS 0 RDA 0	0	0	0	0	0	0	1	0	0	3	0	0	0	0	0
	v	v	v	0	0	0		0	0	5	v	0	U	0	U

Ap	pend	ix F.	50	theo	retic	ally	releva	ant c	oded	segn	nents	(con	tinue	ed)		
															251	L
RDL	. 2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
RDT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RDI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RRS	5 1	0	0	0	0	1	0	1	2	0	1	0	0	2	0	2
RR/	1	0	0	0	0	13	0	4	4	0	7	0	0	0	0	5
RRI	. 5	0	0	0	0	6	0	1	2	0	0	0	0	0	0	2
RR	0 7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RRI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SE	45	14	25	32	10	39	48	14	29	10	49	33	29	57	7	9
KS	65	21	73	53	38	58	61	72	153	41	103	28	16	60	29	28
М	42	21	19	12	3	11	42	53	11	6	6	10	30	96	4	11
ST	7	20	63	7	11	8	30	2	5	6	22	45	45	9	2	2
TC	9	4	10	8	3	17	35	18	11	4	30	2	12	22	8	9
TR	114	26	88	48	11	50	86	102	31	11	78	17	49	151	31	40
то	Г582	199	485	347	246	582	443	428	471	235	496	244	350	641	278	311

 \bigcirc

O

Appendix G. 22 coded segments relevant to revision

252

S1T1 S1T2 S2T1 S2T2 S3T1 S3T2 S4T1 S4T2 S5T1 S5T2 S6T1 S6T2 S7T1 S7T2 S8T1 S8T2

DI	105	70	117	1.00	00	100	00		60		50	<0		=0		• /
RI	125	72	117	163	83	188	80	44	69	66	58	68	74	79	145	56
SP	6	0	1	0	0	1	22	18	3	11	46	0	14	2	1	8
RIS	0	0	0	0	0	2	0	1	11	0	7	0	0	3	0	9
RIA	10	0	0	0	0	11	0	16	4	0	2	0	0	4	0	12
RIL	23	0	1	0	0	9	0	15	10	0	3	0	0	15	0	27
RIT	2	0	0	0	0	3	0	3	3	0	0	0	0	3	0	3
RID	3	0	0	0	0	2	0	7	13	0	0	0	0	9	0	6
RCS	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	3
RCA	0	0	0	0	0	5	0	0	1	0	0	0	0	0	0	1
RCL	2	0	0	0	0	5	0	0	4	0	0	0	0	10	0	1
RCT	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0
RCD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RDS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RDA	0	0	0	0	0	0	0	1	0	0	3	0	0	0	0	0
RDL	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
RDT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RDD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RRS	1	0	0	0	0	1	0	1	2	0	1	0	0	2	0	2
RRA	1	0	0	0	0	13	0	4	4	0	7	0	0	0	0	5
RRL	5	0	0	0	0	6	0	1	2	0	0	0	0	0	0	2
RRT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RRD		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Ū	5	0	5	v	Ū	Ū	Ū	Ŭ	Ū	Ū	U	0	v	Ū	v
тот	180	72	119	163	83	249	102	112	127	77	127	68	88	132	146	136

Appendix H. One-way Analysis of Variance for Revisions From Revisers' Own Input Compared with Revisions From All Other Sources³

Between Subjects

 \square

SOURCE	<u>SS</u>	DF	M\$	<u>F</u>	<u>P</u>
Time	361.000	1	361.000	0.257	0.639
Module	1190.250	1	1190.250	0.846	0.410
Time*Module	144.000	1	144.000	0.102	0.765
Error	5626.500	4	1406.625		

254 Appendix I. One-way Analysis of Variance for Revisions From Revisers' Own Input Compared with Revisions From All Other Sources

Within subjects

SOURCE	SS	DF	MS	F	P
Own-Other	28730.250	1	28730.250	13.852	0.020 *
Own-Other*Time	225.000	1	225.000	0.108	0.758
Own-Other*Module	2756.250	1	2756.250	1.329	0.313
Own-Other*Time*Mo	dule256.000	1	256.000	0.123	0.743
Error	8296.500	4	2074.125		

255 Appendix J. ANOVA for Subject Matter Expert as Source Compared with Audience Expert as Source

Revisions based on feedback data as percentage of all revisions

	Subject matter expert (SME)	Audience expert (TPE)
Mean %	4.415	8.664

SOURCE	SS	DF	MS		<u>P</u>
SME-TPE	72.195	1	72.195	8.351	0.045 *
SME-TPE*Time	32.704	1	32.704	3.783	0.124
SME-TPE*Module	60.579	1	60.579	7.00	0.057
SME-11 E Module	00.379	1	00.379	7.00	0.037
SME-TPE*Time*Module	25.063	1	25.063	2.899	0.164
Error	34.580	4	8.645		

256 Appendix K. ANOVA for Subject Matter Expert as Source Compared with Learner Comments as Source

Revisions based on feedback data as percentage of all revisions

	Subject matter expert (SI	ME) Learner comments
Mean %	4.415	12.254

SOURCE	SS	DF	MS	F	P
SME-					
_		_			
Learner	245.776	1	245.776	11.438	0.028 *
SME-Learner*Time	55.261	1	55.261	2.572	0.184
SME-Learner*Module	9.107	1	9.107	0.424	0.551
SML-Leamer Module	9.107	1	9.107	0.424	0.551
SME-Learner*Time*Modul	le 46.250	1	46.250	2.152	0.216
Error	85.952	4	21.488		

257 Appendix L. ANOVA for Subject Matter Expert as Source Compared with Test Scores as Source

Revisions based on feedback data as percentage of all revisions

	Subject matter expert (SME)	Test Scores
Mean %	4.415	1.784

SOURCE	SS	DF_	MS	<u> </u>	P
SME-Test scores	27.691	1	27.691	9.059	0.040 *
SME-Test scores*Time	1.329	1	1.329	0.435	0.546
SME-Test scores*Module	38.835	1	38.835	12.704	0.023 *
SME-Test scores*Time*Mod	ule 0.545	1	0.545	0.178	0.695
		•	0.515	0.170	0.075
Error	12.228	4	3.057		

Appendix M. ANOVA for Audience Expert as Source Compared with Test Scores as Source

Revisions based on feedback data as percentage of all revisions

	Target population expert (TPE)	Test Scores
Mean %	8.664	1.784

SOURCE	SS	DF	MS		<u>P</u>
TPE-Test scores	189.310	1	89.310	15.186	0.018 *
TPE-Test scores*Time	20.848	1	20.848	1.672	0.266
	20.040	•	20.040	1.072	0.200
TPE-Test scores*Module	2.407	1	2.407	0.193	0.683
TPE-Test scores*Time*Mod	lule32.999	1	32.999	2.647	0.179
Error	49.865	4	12.466		

259 Appendix N. ANOVA for Learner Comments as Source Compared with Test Scores as Source

Revisions based on feedback data as percentage of all revisions

	Learner Comments	Test scores
Means:%	12.254	1.784

SOURCE	SS	DF	MS	<u>F</u>	_ <u>P</u>
Learner-Test scores	438.463	1	438.463	20.957	0.010 *
Learner-Test scores*Time	39.45 1	1	39.451	1.886	0.242
Learner-Test scores*Module	10.330	1	10.330	0.494	0.521
Learner-Test scores*Time*M	odule	36.754	1	36.754	1.757
	0.256				
ERROR	83.688	4	20.922		

Appendix O, ANOVA for Learner Comments as Source Compared with Data From More Than One Source

Revisions based on feedback data as percentage of all revisions

	Learner comments	Data from more than one source
Mean%	12.254	3.773

SOURCE	SS	DF_	MS	F	P
Learner-Data	287.684	1	287.684	16.887	0.015 *
Learner-Data*Time	37.500	1	37.500	2.201	0.212
Learner-Data' Time	37.500	1	57.500	2.201	0.212
Learner-Data*Module	0.425	1	0.425	0.025	0.882
Learner-Data*Time*Mod	ule 45.775	1	45.775	2.687	0.177
ERROR	68.142	4	17.035		

Appendix P. ANOVA for Subject Matter Expert as Source Compared

With Test Scores as Source

C

 \Box

261

Between subjects

SOURCE	SS	DF	MS	<u> </u>	<u>P</u>
Time	2.968	1	2.968	0.567	0.493
	2.908	1	2.900	0.507	0.495
Module	55.969	1	55.969	10.694	0.031 *
Time*Module	2.334	1	2.334	0.446	0.541
F	20.026	4	5 00 4		
Error	20.936	4	5.234		

Appendix Q. ANOVA of Revisers' Use of Personal Input

 \square

C

262

DF-SQUAR	ES DF	MEAN-SQUARE	F-RATIO	P
60.062	1	60.062	0.026	0.876
1105.563	1	1105.563	0.478	0.509
1072.563	1	1072.563	0.464	0.515
430.563	1	430.563	0.186	0.677
138.062	1	138.062	0.060	0.813
2889.063	1	2889.063	1.250	0.296
edback	2376.56	3 1	2376.563	1.029
0.340				
18484.500	8	2310.563		
	60.062 1105.563 1072.563 430.563 138.062 2889.063 edback 0.340	60.062 1 1105.563 1 1072.563 1 430.563 1 138.062 1 2889.063 1 2889.063 1 edback 2376.563 0.340	60.062 1 60.062 1105.563 1 1105.563 1072.563 1 1072.563 430.563 1 430.563 138.062 1 138.062 2889.063 1 2889.063 adback 2376.563 1 0.340	60.062 1 60.062 0.026 1105.563 1 1105.563 0.478 1072.563 1 1072.563 0.464 430.563 1 430.563 0.186 138.062 1 138.062 0.060 2889.063 1 2889.063 1.250 xdback 2376.563 1 2376.563 0.340

0

.

.

.

.