Chapter 13

INTERPRETING THE ADULT LITERACY SCALES AND LITERACY LEVELS

Irwin Kirsch and Ann Jungeblut, Education Testing Service;
Peter Mosenthal, Syracuse University

The major benefits resulting from statistically derived scales are the enhancement of the comparability of results across groups, age, and time, and the provision of a basis for relating background and attitude variables to performance (Messick, Beaton, and Lord, 1983). But however useful such statistically derived scales may be, a need remains to provide supplementary information aimed at guiding the interpretation of the scales. This chapter outlines chronologically the development of a theory of task characteristics that contribute to task difficulty on the three adult literacy scales. This chapter also explains the reasoning that provided the basis for developing five literacy levels on each scale and the process that led to general descriptions of the task characteristics associated with these literacy levels.

13.1 BACKGROUND

Historically, standardized objective tests have provided a means of comparing the performance of one individual to that of a relevant group or groups. A robust theory of mental tests and measurement, including statistical theory and procedures, was developed over the years to facilitate appropriate test interpretation and use. Such techniques have served well in both individual assessments and in large scale surveys. The shift of interest to criterion-referenced testing has yielded more information on task difficulty and the percentages of people able to perform certain tasks. What has been missing is a means of looking at the interaction between task characteristics and people's performance in order to enhance both the interpretability and utility of test results, as well as decisions and actions based on test scores.

The display of both people and tasks along a common scale invites the question of whether tasks receiving similar score values share certain characteristics. This display also raises the issue of to what extent these characteristics (as well as the response consistencies of individuals) differ systematically from one end of the scale to the other. Some important benefits that derive from the systematic exploration of these issues are that they:

- increase understanding of variables that contribute to task difficulty;
- enhance the ability to generate new tasks that more fully represent the domain(s) being assessed;
- establish a context in which one can define the domain boundaries, that is, enhance score meaning; and
- strengthen the links among testing, research, practice, and policy.

Collectively, these benefits contribute to an improved theoretical framework that systematically helps to account for consistency in task responses. Rather than treating the task responses as a conglomeration of specifics, these response consistencies are typically summarized in the form of scores or sub-scores. Although discrete behaviors and isolated observations may be of interest, in terms of measurement validity they are far less meaningful and dependable than response consistencies (Messick, 1989).

The purpose of this chapter is to trace the evolution of the theoretical framework used to construct, interpret, and report large-scale literacy survey data. Prior to the 1992 National Adult Literacy Survey, two other assessments were conducted that used similar methods—the 1985 young adult literacy assessment and the 1990 survey of the literacy of job-seekers served by the U.S. Department of Labor. Together, these three surveys have employed a common definition of literacy, thereby contributing to the evolution of a rich theoretical framework for literacy assessment. The 1992 survey included each of the 1985 literacy blocks in the 1992 cognitive instrument. This chapter describes each study in terms of its contributions to the expanding theoretical understanding of literacy, along with its practical application to literacy measurement, including the development of five proficiency levels used to interpret and report the 1990 and 1992 survey results. The issues and empirical evidence presented address various aspects of validity.

13.2 THE 1985 YOUNG ADULT LITERACY ASSESSMENT

The 1985 young adult literacy assessment was funded with a Federal grant under the National Assessment of Educational Progress (NAEP) program. This survey was designed to assess the literacy skills of young adults 21 to 25 years old. The deliberations of the expert panels that oversaw the development of the young adult literacy assessment led to the adoption of the following definition of literacy: *Using printed and written information to function in society, to achieve one's goals, and to develop one's knowledge and potential* (Kirsch and Jungeblut, 1986a). Reflecting this definition, the organizing theoretical framework that evolved for task development in this study was a multidimensional approach to literacy—that is, tasks were developed to cover the three distinct areas of prose, document, and quantitative literacy.

Literacy tasks for the young adult assessment were based on the printed *material* to be read and the *purpose* which the reader brought to the material. "Material" refers to the linguistic form in which the information is displayed. Twelve categories of material were identified: sign/label, directions, memo/letter, form, table, graph, prose, index/reference, notice, schematic or diagram, advertisement, and bill/invoice. "Purpose" refers to why the reader engages in the task, or what information the reader is seeking. The reader's purpose influences both the strategies and cognitive operations in which the reader engages while completing the task. Five categories of purpose, reflecting various levels of processing were identified:

knowledge, evaluation, specific information, social interaction, and application. The five categories were defined as follows:

- Knowledge—reading to integrate information, to remember sets of facts for later use, or to go beyond information given;
- Evaluation—comparing and contrasting points of view or using printed information to make a reasoned judgment;
- Specific information—-locating a specific fact to satisfy a particular need, such as looking up a fact in a reference book;
- Social interaction—organizing and sequencing information to communicate to another person or group, such as preparing a memo, or writing a letter; and
- Application—following instructions to construct, make, or repair something, doing simple calculations, or providing simple facts to complete forms.

Crossing the twelve identified materials with the five identified uses resulted in the matrix shown below in Table 13-1, taken from Kirsch and Jungeblut (1986b). The dots in the table indicate cells for which literacy tasks were developed and organized into blocks for administration in the 1985 assessment. In developing tasks, primary emphasis was placed on representing the broad range of literacy behaviors people frequently encounter in occupational, social, and educational settings (Guthrie, Siefert, and Kirsch, 1986). It was felt that simulations of the skills in context rather than traditional multiple-choice questions would provide a more ecologically valid and useful assessment of literacy competencies. Efforts were undertaken to create assessment materials that would address these concerns. For example, the 1985 assessment printed a 4-page newspaper containing a selection of articles that had appeared in national newspapers. Respondents were asked to summarize arguments from an editorial, to locate specific information in a news story, and to look up information in a TV listing or a classified page.

Information on the additional considerations underlying the development of the new literacy tasks for the 1992 survey can be found in Campbell, Kirsch, and Kolstad (1992) *Assessing Literacy: The Framework for the National Adult Literacy Survey* and in Chapter 4 of this report.

Table 13-1. Matrix of materials and uses for adult literacy tasks

Materials	Uses				
	Knowledge	Evaluation	Specific information	Social interaction	Application
Sign/Label			✓		✓
Directions					✓
Memo/Letter				✓	
Form		✓	✓		✓
Table		✓	✓	✓	✓
Graph		✓	✓		
Prose	✓	✓	✓	✓	
Index/Reference			✓		
Notice		✓	✓		✓
Schematic or Diagram			✓		✓
Advertisement			✓		✓
Bill/Invoice			✓		

The intersection of the linguistic form in which information is displayed (materials) and the type of information needed or sought (use) in this table not only provided the operational definition of a literacy task but also determined the information-processing demands required for successful performance. The approach to literacy task development guiding the study led to the anticipation of more than a single dimension of literacy. Although the number of cells was too small to allow finding a separate dimension for each of the filled cells in Table 13-1, the designers of the 1985 assessment explored a number of plausible alternative organizing structures on an *a priori* basis before the 1985 data were available for analysis. It is the interaction of the materials and purposes that was expected to define task difficulty and, thus, the placement of tasks on the literacy scales.

13.2.1 Dimensionality of Literacy Skills

Prior to the 1985 assessment, there had been a marked tendency to describe literacy in terms of the ability to perform successfully a series of concrete tasks, e.g., to complete an application for a driver's license, to comprehend the warning on a container of poison, and to interpret familiar street signs (Murphy, 1973). For the most part, success was summed across such diverse tasks and an arbitrary cutting point established (e.g., 75 percent correct), below which an individual is classified as "functionally illiterate" (NAEP, 1976). Such an approach, with its lack of an organizing principle and arbitrary cutpoint used, was dismissed for the 1985 assessment, since it would contribute nothing to understanding the process of literacy. Moreover, this approach was in direct conflict with the theoretical framework of the 1985 study (Kirsch and Jungeblut, with others, 1986, p. III-2).

Previous theoretical approaches to the study of literacy had used a variety of organizing principles. According to one traditional approach, literacy skills had been categorized into reading, writing, speaking, listening, and arithmetic or mathematics. Another approach was to organize disparate literacy tasks by the context in which they occur: home, school, or work. Yet another approach was to categorize literacy tasks in terms of the materials or formats in which they occur and to examine the associated types of purposes both within and across materials. As an instance reflecting a similar distinction, the 1972 NAEP reading assessment came to aggregate reading exercises in terms of "themes"-word meanings, visual aids, written directions, references materials, significant facts, main ideas, inferences, and critical reading (Kirsch and Jungeblut, with others, 1986, p. III-3).

The 1985 assessment designers reached a compromise among the various organizing concepts and hypothesized three scales: a prose literacy scale, a document literacy scale, and a quantitative literacy scale. In this way, they were able to acknowledge that the mental processes underlying proficiency with prose texts are probably qualitatively different from those underlying proficiency with documents and that both are distinct from proficiency with quantities that can be embedded in either prose texts or documents (Kirsch and Jungeblut, with others, 1986, p. III-4).

A factor analysis was performed on the 1985 data in order to explore dimensionality. This was done to find evidence in the empirical data to substantiate the three literacy scales. The product-moment correlation coefficients among the tasks, with squared multiple correlations inserted as communality estimates in the main diagonal, were factor analyzed by the method of principal axes. The mean squared multiple correlation was 0.92 (trace = 101.01). An examination of the latent roots revealed three sizable factors followed by several smaller factors (roots = 18.11, 2.89, 2.30, 2.00, 1.94, 1.87, 1.79, 1.68, 1.67, 1.58,…). Following the logic of Cattell's (1966) *scree* test, the breaks in the pattern of latent roots indicated at least three salient factors with the possibility of at least five additional factors. Analysis of parallel random data reinforced the judgment that a three-factor solution was appropriate. However, for exploratory purposes three separate analyses were conducted: one in which eight factors were retained and rotated for interpretation; another in which five factors were retained; and, a third in which, three factors were retained for rotation and interpretation (Kirsch and Jungeblut, with others, 1986, p. III-5).

In each instance, the factors were rotated to simple structure by the varimax procedure and to oblique simple structure by the DAPPER method (Tucker and Finkbeiner, 1981). The DAPPER method was selected specifically to allow the complex literacy tasks to load on more than one factor. Indeed, many of the literacy tasks did so (Kirsch and Jungeblut, with others, 1986, p. III-6). Tasks loading highest on the first and largest factor seemed to rely heavily on prose comprehension; tasks loading highest on the second factor seemed to reflect skill in using documents, while those tasks loading highest on the third factor

required the application of arithmetic operations. The DAPPER method was selected specifically to allow the complex literacy tasks to load on more than one factor.

The intercorrelations among the literacy scales for the total group of 3474 respondents ranged from .49 to .56, thus revealing only a moderate level of association. Indeed, the intercorrelations provide further support for the notion that literacy skills can and should be separated along at least three distinct dimensions—prose, document, and quantitative skills. These important distinctions would be lost if the diverse tasks from the 1985 young adult literacy assessment had been aggregated and reported on a single scale (Kirsch and Jungeblut, with others, 1986, p. III-34).

13.2.2 Difficulty of Literacy Tasks

Since the 1985 young adult literacy assessment was funded under the NAEP program, it tended to adopt many of the survey design and statistical methods used in that program. Beginning with 1983-84 reading assessment of school children, NAEP chose to anchor items representing standard deviation units along the reading proficiency scale. The exemplar items selected discriminated between each pair of standard deviation units in the following way: The NAEP reading proficiency scale was designed to extend from 0 to 500 with a mean of 250 and a standard deviation of 50. Thus, the selected anchor points were 150, 200, 250, 300, and 350 (Beaton and Allen, 1992). The criteria for selecting exemplars at each anchor point were that 80 percent or more of the students at that point (e.g., 250) answered the item correctly, while less than 50 percent of the students at the next lower level (e.g., 200) answered the same item correctly (Kirsch and Jungeblut, with others, 1986, p. III-9). The exemplar items identified through this procedure represented advances in student reading proficiency from one anchor point to the next. A panel of content experts then examined the exemplar items near each anchor point and identified the reading knowledge, skills, and abilities demonstrated by students answering each item correctly. The panelists' descriptions were summarized to characterize performance at each anchor level (Phillips, et al., 1993).

In the 1985 assessment, the three literacy scales were designed to range from 0 to 500, with a mean of 250 and a standard deviation of 50 points.[1] Unlike the 1983-84 NAEP reading assessment, the 1985 young adult literacy assessment had relatively few literacy tasks on the prose and quantitative literacy scales, so it was not feasible to provide meaningful descriptions at identical numerical points (e.g., standard deviation units) on each of the three scales. In addition, one would not expect that on each of the scales, tasks exemplifying important shifts in their cognitive demands would fall at comparable points (Kirsch and Jungeblut, with others, 1986, p. III-9).

[1]The scales used in 1992 were linked to the 1985 scales, so the mean could not be fixed at 250. The IRT models that structure these scales are described in Chapter 9.

To help guide interpretation of the three literacy scales and to attempt to identify factors associated with task difficulty, an item mapping procedure reflecting response probabilities was employed. Following NAEP's anchoring system at the time, tasks were placed on the scale at the point at which a minimum of 80 percent of the young adults at a particular ability level could be expected to complete the tasks successfully (see Chapter 14 for additional information). An additional criterion for selection of exemplar tasks was that approximately 30 percentage points (in terms of IRT response probabilities) separated individuals scoring around the same scale value as the task placement from individuals scoring one standard deviation (i.e., 50 points) lower. In the context of the 1985 young adult literacy assessment, the particular exemplars not only met this anchoring criterion representing performance at various levels of difficulty but also were seen to reflect a combination of increasingly complex skills interpreted to be associated with successful performance as task difficulty increased.

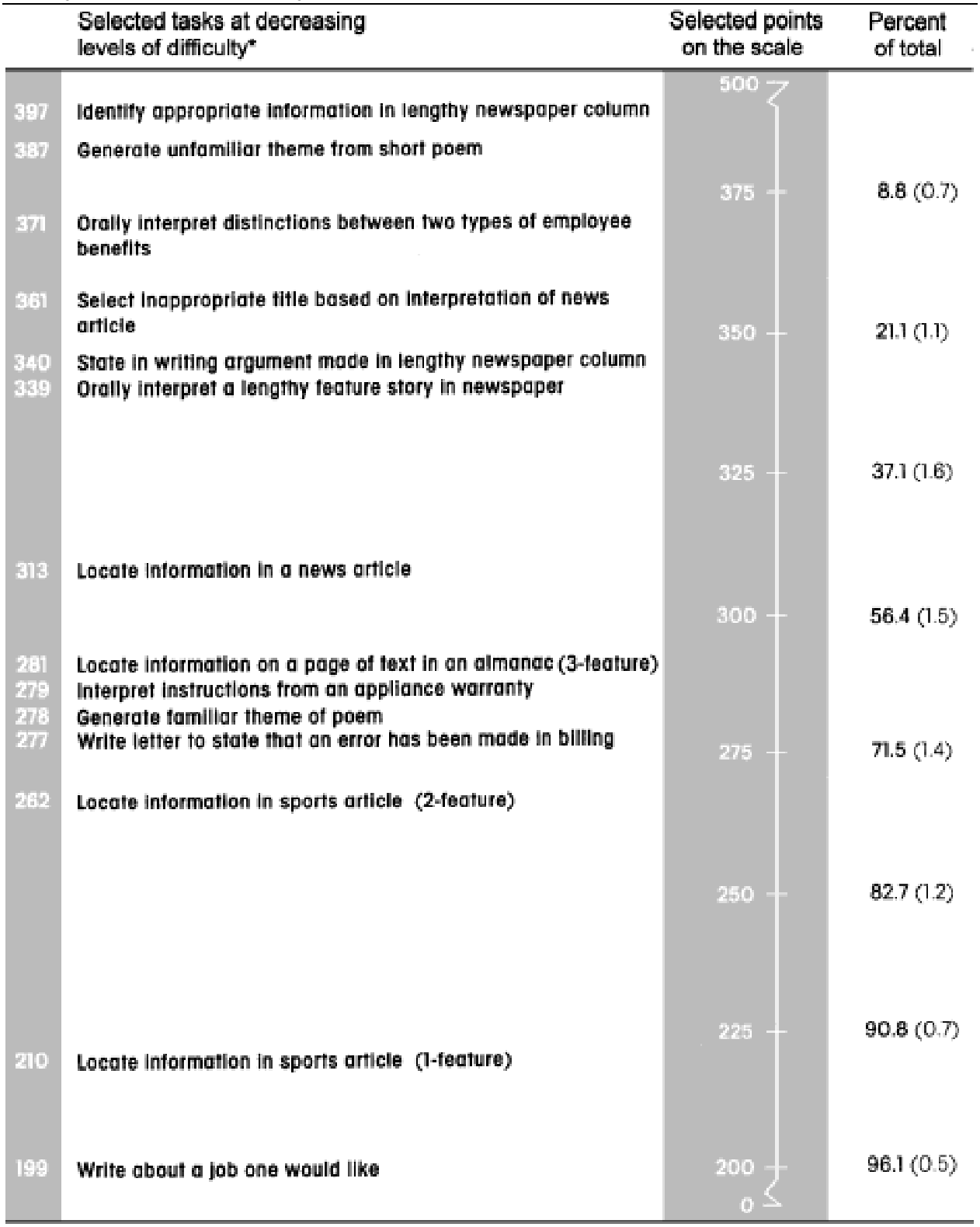
Once tasks were located on the scales, it was possible to summarize survey results by presenting selected exemplar tasks placed around successive points on the three scales along with percentages of people scoring at or above those same points. It was anticipated that such a graphic presentation would help give meaning to the scale and, thereby, increase the interpretability of results for the total group, as well as for subgroups of particular interest.

The next three sections of this chapter describe the mapping of task difficulty, the proficiencies of young adults on the 1985 prose, document, and quantitative scales and the identification of characteristics that underlie task difficulty at various points on the three scales. Task characteristics were identified on the basis of the complexity of the information-processing demands required for successful performance, rather than by features of the text alone, such as vocabulary or sentence length.

13.2.3 Prose Comprehension Scale

Exhibit 13-1 presents information about task difficulty and population performance on the prose literacy scale based on the item map reported from the young adult assessment (Kirsch and Jungeblut, 1986a). It can be readily seen, for example, that 90 percent or more of young adults in America demonstrated proficiency on the tasks extending down from 225 on the prose literacy scale. The tasks associated with points below 225 in the column on the left range from skill in locating one feature of information in a sports article to writing about a job one would like. In the original report, the full figure (not reproduced here) showed results not only for the total population but also for various racial/ethnic groups and for different levels of educational attainment. While each subgroup differed in their mastery at various levels on the scale, the ordering of the task difficulties remained the same across all groups.

Exhibit 13-1. Percentages of adults and selected tasks at or above successive points on the prose literacy scale: Adults 21 to 25 years old, 1985



*Number indicating difficulty level designates that point on the scale at which individuals with that level of proficiency have an 80 percent probability of responding correctly.

The 15 tasks comprising the prose comprehension scale appeared to reflect three qualitatively different aspects of reading comprehension: 1) matching of literal and corresponding information; 2) producing and interpreting text information; and 3) generating a theme or organizing principle from text information. Each of these three aspects contributed to a broad range of difficulty, with significant overlap among the three. Exhibit 13-2 presents information about task difficulty for the three aspects of the prose comprehension scale, again based on an item map reported in the young adult assessment (Kirsch and Jungeblut, 1986a).

Proficiency in Matching Literal Corresponding Information represents a continuum defined by the number of features that readers must identify to match information asked for in a question or directive with explicit or corresponding information in the text. At the simplest end of the continuum, readers match requested information with information in the text on the basis of a single, commonly shared feature. At the middle and upper ends, the match involves several features or several categories of information.

Proficiency in Producing and Interpreting Text requires readers to use background knowledge or textual information. Response at the simplest level involves producing personal background information. At more difficult levels, readers may have to interpret the directive or compare and contrast information.

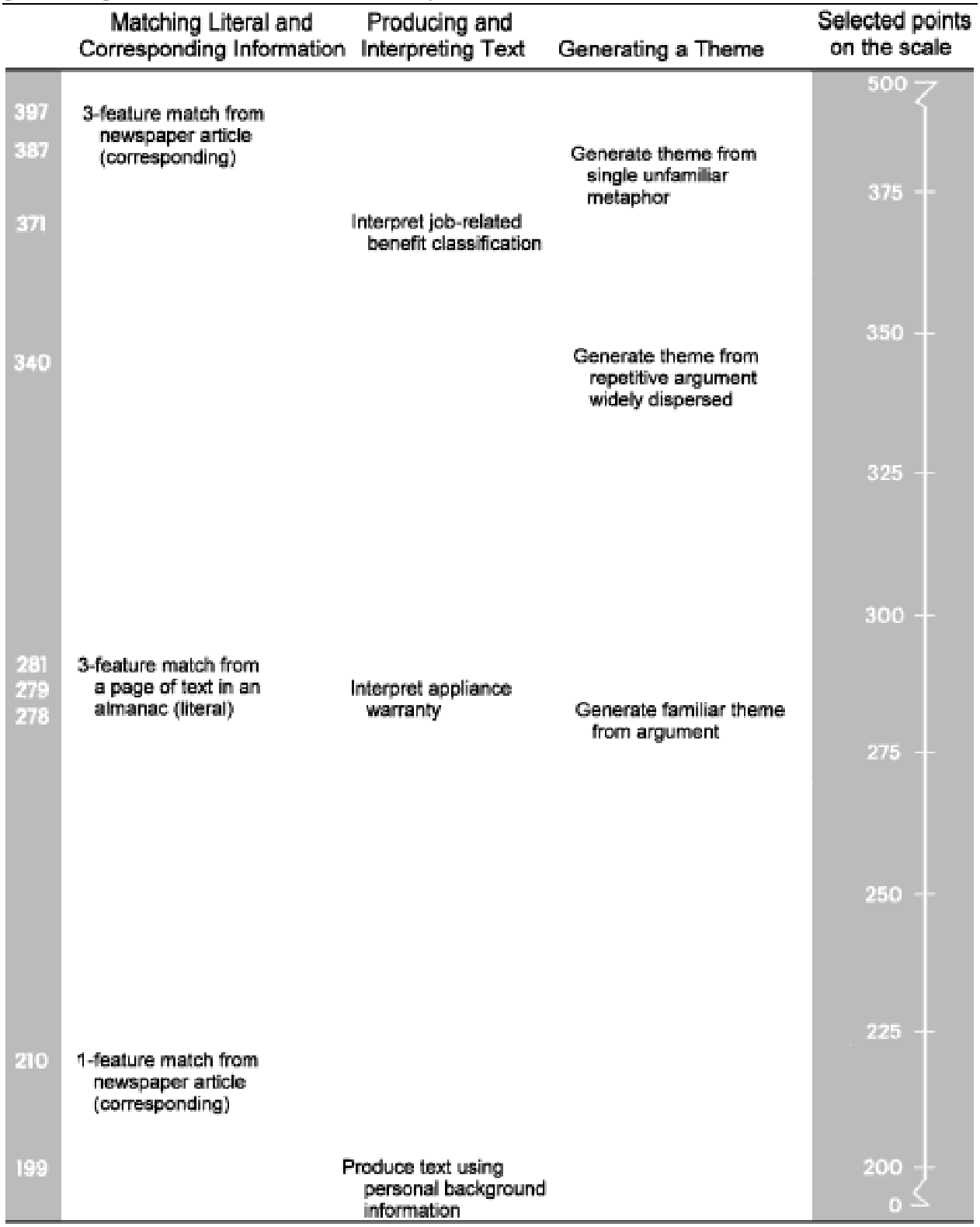
Proficiency in Generating a Theme or Organizing Principle from Text Information requires readers to synthesize information consistent with arguments in the text. At the simplest level, the reader's task is to generate a theme from relatively short text. Generating the theme becomes more difficult as the concept becomes less familiar, or the arguments are less repetitive or more widely separated in lengthy text.

13.2.4 Document Literacy Scale

Exhibit 13-3 reproduces information about task difficulty and population performance on the document literacy scale based on the original item maps from the young adult literacy assessment (Kirsch and Jungeblut, 1986a). Again, it can be seen that 90 percent or more of the total group of young adults surveyed demonstrated proficiency on the tasks extending down from 225 on the document literacy scale. The tasks associated with points below 225 in the column on the left include entering the date on a bank deposit slip, identifying the cost of a particular theater trip from among those given in a notice, entering personal information on a job application form, locating the expiration date on a driver's license, and signing one's name on an image of a Social Security card.

The 43 tasks comprising the 1985 document literacy scale begin with a question or directive. The reader must first identify the important information in the question to be matched to information in a document. Among the most important characteristics associated with task difficulty are: 1) the number of features that readers must identify in a question or directive and match with features of information in a

Exhibit 13-2. Selected tasks and corresponding levels of difficulty* defining the three aspects of the prose comprehension scale: Adults 21 to 25 years old, 1985



*Number indicating difficulty level designates that point on the scale at which individuals with that level of proficiency have an 80 percent probability of responding correctly.

Exhibit 13-3. Percentages of adults and selected tasks at or above successive points on the document literacy scale: Adults 21 to 25 years old, 1985

	Selected tasks at decreasing levels of difficulty*	Selected points on the scale	Percent of total
343	Use bus schedule to select appropriate bus for given departure	500	
320	Use sandpaper chart to locate appropriate grade given specifications	325	37.6 (1.6)
300	Follow directions to travel from one location to another using a map	300	57.2 (1.7)
294	Identify information from graph depicting source of energy and year		
278	Use index from an almanac	275	73.1 (1.2)
262	Locate eligibility from table of employee benefits		
257	Locate gross pay-to-date on pay stub		
255	Complete a check given information on a bill		
249	Locate intersection on street map	250	83.8 (1.0)
221	Enter date on a deposit slip	225	91.0 (0.8)
219	Identify cost of theatre trip from notice		
211	Match items on shopping list to coupons		
196	Enter personal information on job application	200	95.5 (0.5)
192	Locate movie in TV listing in newspaper		
181	Enter caller's number on phone message form		
169	Locate time of meeting on a form	175	98.4 (0.3)
160	Locate expiration date on driver's license		
		150	99.7 (0.1)
110	Sign your name	0	

*Number indicating difficulty level designates that point on the scale at which individuals with that level of proficiency have an 80 percent probability of responding correctly.

document; 2) the degree to which feature information given in the question or directive corresponds to, or is closely identified with, the requested information in the document; and 3) the number of exemplars or representations in the document that have at least one feature in common with those in the question, thereby serving as distractors or plausible correct answers for the reader.

Once a match between a question (or directive) and document information is made, the reader must determine whether the information matched is sufficient. If it is insufficient, the reader must cycle back through the process. This might require the reader to re-identify features in a question or directive or to re-enter the document and to search and locate additional features. Once the reader determines that sufficient information has been matched, the task can be executed by completing the directive.

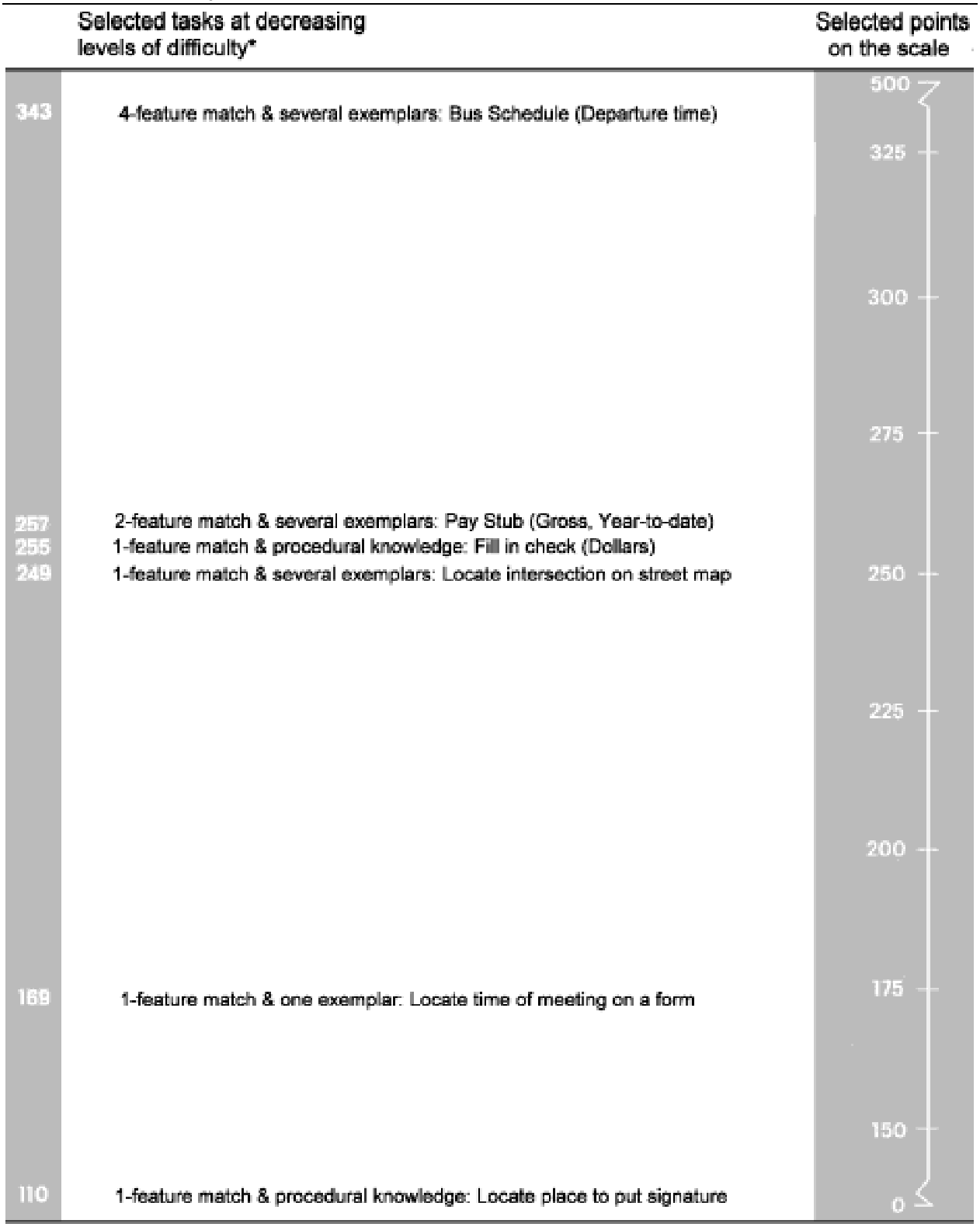
Exhibit 13-4 presents information about task difficulty for the document literacy scale, again based on an item map reported in the young adult literacy assessment (Kirsch and Jungeblut, 1986a). Proficiency in matching document information represents a continuum defined by the number of features that readers must identify in a question or directive. At the simplest end of the continuum, readers match literal information on the basis of a single feature in a document that contains only one exemplar. At the middle and upper ends, the match involves increasing numbers of features. In some cases the matches are literal, while in others the matching is based on varying degrees of correspondence.

Task difficulty increases along with increases in the number of features to be matched, the number of exemplars serving as distractors, and the degree to which information in the question or directive lacks correspondence or identity with the needed information in the document. This aspect of document literacy tasks not only had a dominant role in defining the difficulty of document literacy tasks, but also in defining task difficulty on the prose comprehension scale and, to a lesser extent, on the quantitative scale—that is, matching information in a question or directive with literal or corresponding (synonymous) text information.

13.2.5 Quantitative Literacy Scale

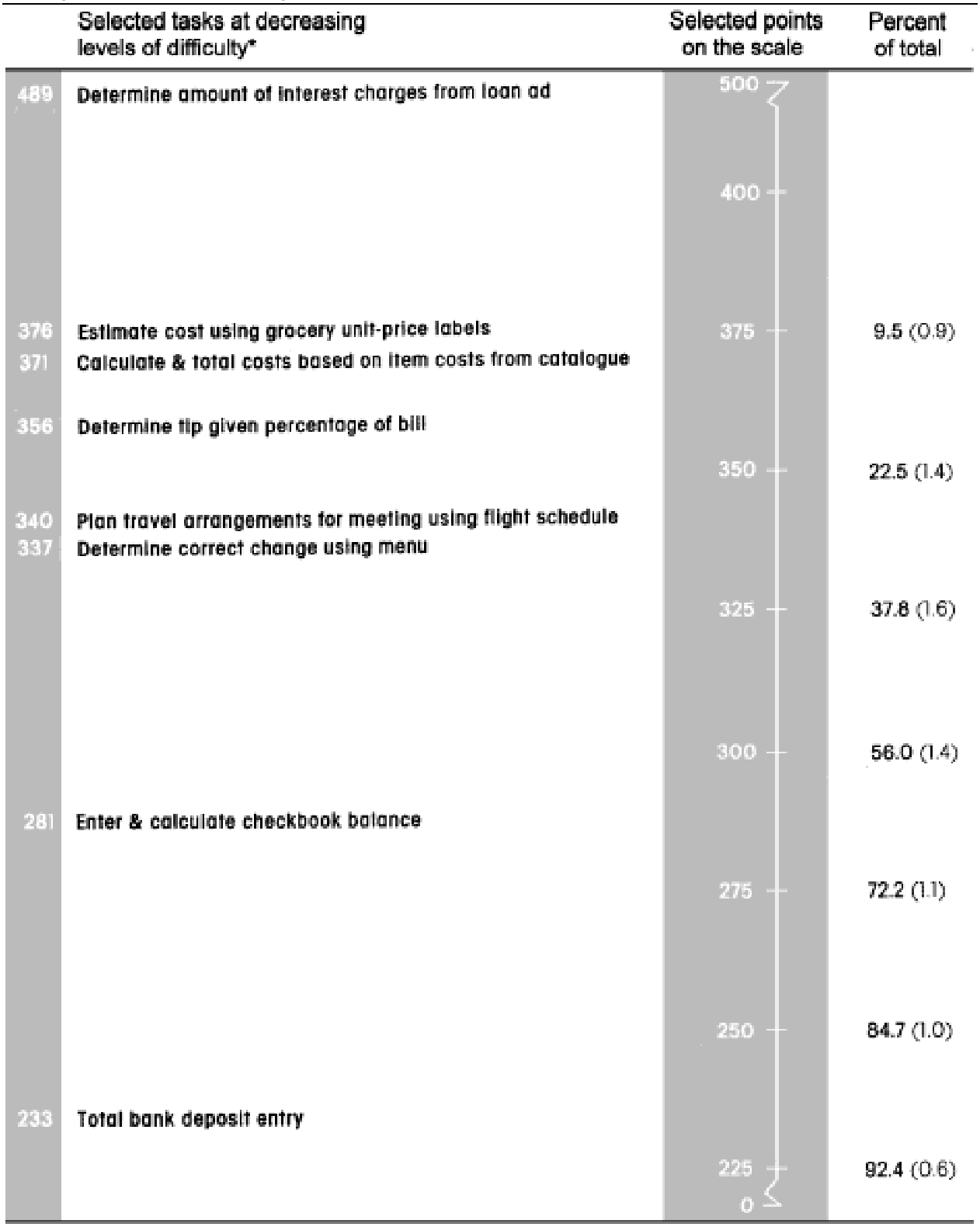
Exhibit 13-5 presents information about task difficulty and population performance, this time on the quantitative literacy scale, based on an original item map from the young adult literacy assessment (Kirsch and Jungeblut, 1986a). Again, it can be seen that 90 percent or more of the total group of young adults surveyed demonstrated proficiency on the tasks extending down from 225 on the quantitative literacy scale. On this scale there were no tasks associated with only an 80 percent success rate for points below 225. Subsequent literacy assessments developed literacy tasks capable of distinguishing using the 80 percent criterion, at the lower end of the scale.

Exhibit 13-4. Selected tasks and corresponding levels of difficulty* defining the document literacy scale: Adults 21 to 25 years old, 1985



*Number indicating difficulty level designates that point on the scale at which individuals with that level of proficiency have an 80 percent probability of responding correctly.

Exhibit 13-5. Percentages of adults and selected tasks at or above successive points on the quantitative literacy scale: Adults 21 to 25 years old, 1985



*Number indicating difficulty level designates that point on the scale at which individuals with that level of proficiency have an 80 percent probability of responding correctly.

The 15 tasks comprising the 1985 quantitative literacy scale appeared to reflect the ability to use mathematical operations such as addition, subtraction, multiplication, or division, either singly or in combination, to solve problems variously embedded in printed material.

Factors associated with task difficulty and performance on the quantitative scale appeared to be the type of arithmetic operation (addition, subtraction, multiplication, and division) required for a correct answer, the number or combination of operations needed, and the extent to which the specification of the operations are embedded in textual material. At the simplest end of the continuum, readers carry out a single, specified operation on numbers that appear in convenient places on the document. At the next level of difficulty, tasks require a single operation, but they also require that the reader enter the appropriate information from the question or directive onto the document before the operation can be completed. At a more difficult level, tasks require either two sequential operations or the application of a single, higher level operation (multiplication or division). At the upper end, the tasks require disembedding the appropriate features of a problem (in the presence of distractors) and carrying out a sequence of operations.

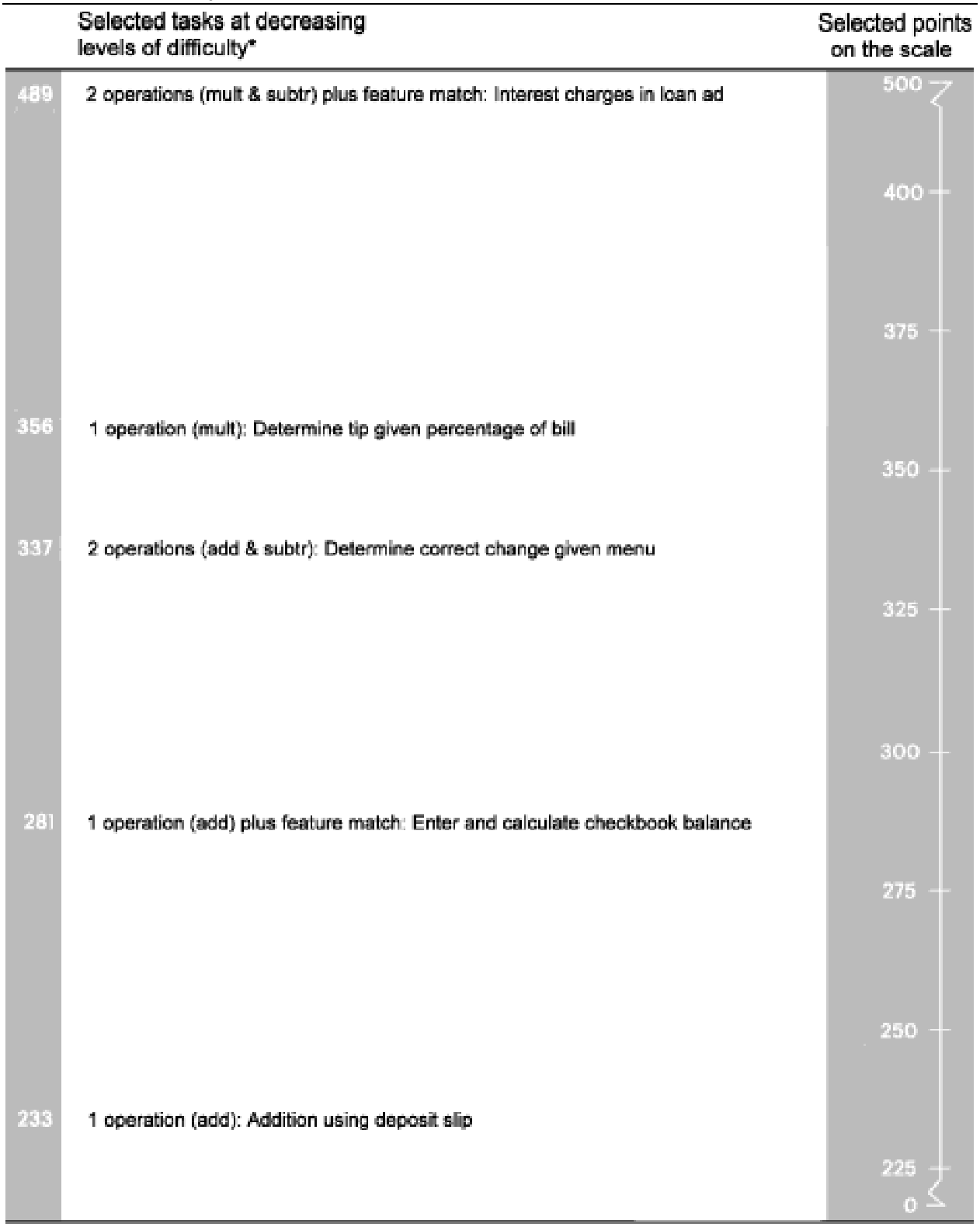
Task difficulty on the quantitative scale was associated not only with the type of operation but also with the number of operations required and the degree to which the problem is embedded in printed material. Exhibit 13-6 presents information about task difficulty for the quantitative literacy scale, based again on an item map reported in the 1985 young adult literacy assessment (Kirsch and Jungeblut, 1986a).

13.3 ENHANCING UNDERSTANDING OF TASK DIFFICULTY

Following publication of the final report from the 1985 young adult literacy assessment, Kirsch and Mosenthal undertook a secondary analysis of the 1985 assessment data in order to extend the understanding of factors associated with document task difficulty—that is, the location of tasks along the document literacy scale (Kirsch and Mosenthal, 1990). They applied an extensive grammar to the 37 different stimulus documents (representing nine categories, e.g., tables, graphs, charts, and forms) as well as to the questions or directives of the 61 specific document literacy tasks developed for the young adult assessment. Their intent was to describe the structure and content of diverse printed materials.

The labor-, knowledge-, and time-intensive grammar devised by Kirsch and Mosenthal relied on semantic-relation categories drawn from other propositional grammars (Mosenthal, 1985). Their application of the grammar to both stimulus materials and associated task directives revealed structural patterns in the document tasks, patterns that encouraged Kirsch and Mosenthal to hypothesize about the variables underlying performance on the literacy tasks. These hypothesized variables related to the structure and complexity of the document or stimulus material, to the nature of the task (i.e., the structural relations between the document and the question or directive), and to the nature of the processes readers engaged in to relate information in the question or directive to information in the document. They

Exhibit 13-6. Selected tasks and corresponding levels of difficulty* defining the quantitative literacy scale: Adults 21 to 25 years old, 1985



*Number indicating difficulty level designates that point on the scale at which individuals with that level of proficiency have an 80 percent probability of responding correctly.

identified six document variables, four task variables, and three process variables they believed would contribute to difficulty in processing documents. Their initial analyses, using percent correct statistics (rather than response probabilities), identified 12 of the 13 hypothesized variables within their three major categories of variables as significantly influencing the demonstrated difficulty of document literacy tasks for young adults.

Kirsch and Mosenthal further reduced these 12 variables down to a set of eight by omitting those variables with a zero-order correlation of less than .30 with percent correct. Subsequent regression analysis of these eight variables showed that there were five variables that contributed significantly to variance in the percentage correct scores and were consistent (i.e., generalizable) across both racial/ethnic groups and levels of educational attainment (Kirsch and Mosenthal, 1990, Table 2).

Of the final set of five significant variables, two were process variables: degree of correspondence and type of information (Mosenthal and Kirsch, 1991a). *Degree of correspondence* refers to the first stage of document processing, in which a reader must match information given in a question with corresponding information in a document, and varies from easiest (literal or synonymous correspondence) to hardest (correspondence arrived at via high, text-based inference or using special prior knowledge) (Kirsch and Mosenthal, 1990, p. 19). While the *degree of correspondence* variable deals with the correspondence between the information given in the question and the information contained in the document, *type of information* focuses primarily on the requested information. More specifically, type of information refers to how the reader obtains that information, by locating, identifying, generating, or synthesizing requested information based on various "nodes" of a document's information hierarchy. Document processing becomes more difficult as: (a) the reader must generate inferences or use prior knowledge to relate the request to the document, or (b) the reader must relate information across different nodes to arrive at a response (Kirsch and Mosenthal, 1990, p. 20).

Of the final set of five significant variables, two others were task variables (the number of organizing categories and the number of specifics that needed to be processed to complete a task successfully). These variables represent the two types of structural relations between a question or directive and the document and quantify aspects of the amount of information that the reader must process. The *number of organizing categories* consists of the quantity of labels that serve to summarize or synthesize specific data or entries in a document. The *number of specifics* deals with the number of entries or pieces of information the reader must process in order to respond correctly to the task.

Of the final set of five significant variables, the remaining variable was a document variable (the number of specifics), involving the length and complexity of the document itself. The *number of specifics*

was a measure of the length and amount of material that was contained in the document. As the number of specifics increased, so did the difficulty of the document.

One additional process variable was notable, though it was not among the final set of five significant variables—plausibility of distractors. This variable refers to the situation where information in a text or document meets some but not all of the conditions required in a question or directive to provide a correct response. Despite the fact that this variable failed to reach significance for the various racial/ethnic and educational attainment groups, it had one of the highest zero-order correlations with percent correct scores among various subgroups of interest.

With the exception of young adults reporting zero to eight years of education, the variance in task difficulty accounted for by the subset of five significant variables ranged from 89 percent (for the total group assessed and for White young adults) to 81 percent (for both black young adults and those who dropped out of high school before earning a diploma). Some 56 percent of the variance in percentage correct scores was accounted for in the group of individuals reporting zero to eight years of schooling. In general, then, the results yielded strong empirical validity evidence for both the evolving theory and for document score interpretation.

This study provided not only a theoretically-based model of performance but also an applied means of predicting task difficulty, along with identified cognitive characteristics for the set of literacy tasks included in the 1985 young adult literacy assessment (Mosenthal and Kirsch, 1991). Although not reported in the literature because of the relatively small number of tasks, the same approach was also applied to both the prose and quantitative scales. Given their enhanced understanding of a set of variables that seem to underlie successful performance on document literacy tasks, Kirsch and Mosenthal used this knowledge to devise specifications for developing new tasks targeted to specific degrees of difficulty along the literacy scales.

The original coding scheme was useful in the design and development of new tasks written for the 1990 survey of the literacy skills of Job Training Partnership Act (JTPA) and Employment Service/Unemployment Insurance (ES/UI) program participants. Additional evidence for the validity of the theory and for score interpretation rests on the success of this task development work.

13.4 THE 1990 SURVEY OF THE LITERACY OF JOB-SEEKERS

The 1990 survey of the literacy skills of job-seekers served by the U.S. Department of Labor (Kirsch and Jungeblut, 1992) capitalized on the results of the 1985 young adult literacy assessment, as well as on the secondary data analyses conducted by Kirsch and Mosenthal. The definition of literacy, the three literacy scales, and the expanded theoretical framework all contributed to the 1990 survey. While the earlier developments provided an important initial step, several of the variables required using the complex and

labor intensive grammar. Through a series of revisions and enhancements, a set of variables was identified that eliminated the need for using the grammar and greatly improved the utility of the coding procedures for others interested in the area of literacy. Before discussing the 1990 survey, it will be useful to briefly describe how the current set of variables for prose, document, and quantitative tasks evolved from the original research.

The process variables Kirsch and Mosenthal identified as important through their secondary analysis included *degree of correspondence*, *type of information*, and *plausibility of distractors.* *Plausibility of distractors* was the only variable to remain unchanged throughout this process. A new process variable, *type of match*, was developed by merging the original *type of information* variable with the *degree of correspondence* variable. A third process variable, *type of information*, was also added to indicate the degree of abstractness of the information requested in a question or directive.

Finally, there was some concern that the process variables identified and the associated variance being accounted for might possibly reflect simply the notion of "readability," which has a long history in theoretical and applied research. To address this issue, an estimate of readability was devised from the grammar for use with document stimuli (Mosenthal and Kirsch , 1998), and from the Fry (1977) formula for use with prose stimulus materials.

The following sections evaluate the utility of the current framework as it relates to the creation of new tasks, the understanding of the variables contributing to task difficulty, and the enhancement of score meaning in the 1990 survey of the literacy skills of job-seekers.

13.4.1 Prose Literacy

An important area of literacy is the knowledge and skills needed to understand and use information organized in sentence and paragraph formats. Given the range of text types organized in such formats, the 1990 job-seeker assessment used prose materials that were primarily expository (i.e., materials which describe one or more states or actions) since such materials constitute much of the prose that adults read (Kirsch and Jungeblut, 1986a; Kirsch et al., 1992). In addition, some narrative texts and poetry were included. The prose materials were drawn from newspapers, magazines, books, brochures, and pamphlets, and were reprinted in their entirety, using the typography and layout of the original source. As a result, the materials varied widely in length, in density of information, and in the use of structural or organizational aids, such as section or paragraph headings, italic or bold face type, and bullets.

13.4.1.1 Prose variables

Prose tasks involve the problem of first identifying *given* and *requested* information (Fisher, 1981; Clark and Haviland, 1977; Mosenthal and Kirsch, 1991). *Given* information is that which is known and assumed to be true based on the way a question or directive is stated. *Requested* information in a question or

directive is that which is being sought. To illustrate this, consider the question, “In the past five years, how many times has Susan Butcher won the Iditarod Sled Dog Race?” The *given* information in this instance is “In the past five years, Susan Butcher won the Iditarod Sled Dog Race one or more times.” The *requested* information of this sentence is “*How many times* did Susan Butcher win?” In processing prose, tasks tend to be easy when the requested information is concrete; tasks tend to be more difficult the more abstract the requested information becomes. Hence, a task whose requested information involves a person or thing (e.g., a *who* or *what* question) tends to be easier to process than a task whose requested information asks for a reason, purpose, or cause (e.g., a *why* question) (Mosenthal, 1998).

Another dimension of prose processing requires readers to match information in a question or directive to corresponding information in a text. This involves the strategies of *locating, cycling, integrating,* and *generating* information. *Locating* tasks require the reader to find information in the text based on conditions or features specified in the question or directive. The match may be literal or synonymous, or the reader may need to make an inference in order to perform successfully. *Cycling* tasks require the reader to locate and match one or more features but also require the reader to engage in a series of feature matches to satisfy conditions given in the question. *Integrating* tasks require the reader to compare or contrast two or more pieces of information from the text. In some cases the information can be found in a single paragraph, while in others it appears in different paragraphs or sections. In the *generating* tasks, readers must produce a written response by processing information from the text and also by making text-based inferences or drawing on their own background knowledge. These processes are represented by the variable *type of match* (Mosenthal, 1998).

A third dimension of prose processing involves *plausibility of distractors,* the situation where information in text meets some but not all the conditions of the answer specified in the question or directive. The more conditions that such distracting information shares with a correct answer and the more closely it is positioned to the correct answer, the more difficult the processing becomes (Mosenthal, 1996).

In addition to the three process variables, Kirsch and Mosenthal considered a fourth variable—readability—representing complexity of prose materials (Fry, 1977). It was included not only to determine the extent to which it accounted for task difficulty, but also to provide another descriptor commonly found in the research literature on prose processing (Mosenthal, 1998).

Kirsch and Mosenthal evaluated each of the 45 prose tasks included in the 1990 job-seeker assessment in terms of these four variables. They devised a coding scheme and applied it to each of the prose literacy tasks. The variables *type of information* and *plausibility of distractors* range from one (easiest) to five (most difficult). The coding scheme they applied to *type of match* was additive and ranged from one to a possible 20, although the type of match variable for actual tasks in the 1990 survey ranged

from only one to eight. Based on this scheme, several example tasks are presented in the next section that highlight the range of task complexity required for successful performance along this dimension of literacy.

13.4.1.2 Examples of prose literacy tasks

One of the easiest prose tasks (RP80 difficulty value of 210) involved a short newspaper article about a marathon swimmer (Exhibit 13-7). This text reflects an eighth-grade Fry readability level. The directive asks the reader to "underline the sentence that tells what Ms. Chanin ate during the swim." To complete this directive, readers have to recognize that the requested information is a thing (i.e., food). This prose task received a code of 1 for the *type of information* process variable. In identifying the requested information, readers must make a synonymous match between "ate" in the directive and "banana and honey sandwiches, hot chocolate, lots of water, and granola bars" in the text. This task received a code of 1 for the *type of match* process variable. Note that, since there is no other mention of food in the text, there are no plausible distractors for requested information. This task received a code of 1 for the *plausibility of distractors* process variable.

Exhibit 13-7. Example of text for relatively easy prose literacy tasks

Swimmer completes Manhattan marathon

The Associated Press

NEW YORK-University of Maryland senior Stacy Chanin on Wednesday became the first person to swim three 28-mile laps around Manhattan.

Chanin, 23, of Virginia, climbed out the East River at 96th Street at 9:30 p.m. She began the swim at noon on Tuesday.

A spokesman for the swimmer, Roy Brunett, said Chanin had kept up her strength with "banana and honey" sandwiches, hot chocolate, lots of water and granola bars."

Chanin has twice circled Manhattan before and trained for the new feat by swimming about 28.4 miles a week. The Yonkers native has competed as a swimmer since she was 15 and hoped to persuade Olympic authorities to add a long-distance swimming event.

The Leukemia Society of America solicited pledges for each mile she swam.

In July 1983, Julie Ridge became the first person to swim around Manhattan twice. With her three laps, Chanin came up just short of Diana Nyad's distance record, set on a Florida-to-Cuba swim.

Find the article "Swimmer completes Manhattan marathon" on page 2 or the newspaper provided and answer the following questions.

11. Underline the sentence that tells what Ms. Chanin ate during the swim.

12. At what age did Chanin begin swimming competitively? _____

A second task involving this text has an RP80 scale value of 250. This task includes the question, "At what age did Chanin begin swimming competitively?" To answer this question, readers must recognize that the requested information is an amount (i.e., age). This task received a code of 2 for the *type of information* process variable (because an amount in this task is more abstract than a thing in the previous task). To identify the requested information, readers have to make a synonymous match. Having

made this match, readers then must find the answer "15" in the sentence "The Yonkers native has competed as a swimmer since she was 15 and Y." This task received a 3 for the *type of match* process variable (because a cycle of matching was required). What also makes this task somewhat difficult is the fact that there is a distractor for the requested information that appears earlier in the text; this distractor includes the information that the swimmer's current age is 23. This task received a 4 for the *plausibility of distractors* process variable (because distractors appear for both given and requested information, but not in the same paragraph as the answer).

A task receiving an RP80 scale value of 247 involved a rather lengthy article on parenting written by Dr. Spock (Exhibit 13-8). While this article only represented an eighth-grade Fry readability level, it consists of a relatively long passage without any organizational aids (a challenge not measured by the Fry approach). One question asked the reader to identify one alternative to the use of physical punishment. The reader could match the phrase "alternative to the use of physical punishment" to the phrase "other punishments parents can use." The text then lists a number of alternatives recommended by Dr. Spock. *Type of match* received a score of 2 (because cycling was required) and *plausibility of distractors* also received a score of 2 (because similar information appears somewhere in the text, but not nearby), while *type of information* received as score of 3 (because an alternative is more abstract than a thing or an amount).

A somewhat more difficult task based on the same text (RP80 score of 283) requires the reader to "list the two reasons given by the author why physical punishment is still widely accepted as a way to teach children right and wrong." This task can be answered by locating the place in the text that begins, "I think there are two reasons for this. The first is .... The second reason is ...." *Type of match* received a score of 3 (an extra point was added to a 2 for cycling by the need to identify the antecedent of the pronoun "this"). *Plausibility of distractors* received a score of 2 (because similar information appears somewhere in the text, but not nearby). *Type of information* received an abstractness score of 4 (because a reason is more abstract than an alternative).

The most difficult task involving this text (an RP80 score of 311) directs the reader to "identify and list two reasons that Dr. Spock offers for not using physical punishment." While numerous statements throughout the article help satisfy the directive, much of the text deals with related concerns rather than direct summary statements. As a result, the reasons for not using physical punishment are embedded throughout the text and are not literally stated following a semantic cue such as "Two good reasons for not using physical punishment are Y." This task was coded 3 for *type of match* (because it requires synthesizing features across the document). In addition, distracting information is more closely tied to words or phrases containing the necessary information for responding correctly. This task was coded 3

Exhibit 13-8. Example of text for moderately difficult prose literacy tasks

PARENTING
BY BENJAMIN SPOCK, M.D.

Have You Ever Wanted To Strike Your Child?

Don't do it! Dr. Spock believes that physical discipline can cause lasting resentment in a sensitive child and may make a naughty child a real behavior problem.

Almost all parents with whom I've ever discussed the issue of physical punishment acknowledge that they've had a strong impulse to spank their children at one time or another, whether they believed in doing it or not: for instance, when a small child breaks a valuable object she has been told not to touch, or when a somewhat older child of six or seven runs into the street and a car just misses hitting him, or when an eleven-year-old is caught stealing and then brazenly tries to lie her way out of it. And it's the rare parent who has *never* given in to the impulse to slap or spank.

Parents tend to punish their children the same way their own parents punished them — whether it's by spanking or scolding or reasoning or withholding privileges. In this way patterns of discipline — both good and bad — are passed from one generation to the next.

Why is it that physical punishment, whether used occasionally or frequently, is still widely accepted as a way of teaching children what is right and what is wrong? I think there are two reasons for this. The first is the belief that it is simply the correct way of handling certain kinds of misbehavior, such as those I've mentioned earlier. The second reason is even more powerful, and it has to do with the parent's *reaction* to the misbehavior: the wave of anger that sweeps over the parent when a child misbehaves, *especially* when there is an element of defiance in an act or in an attitude. The child's challenge to the parent's authority causes a spasm of panic: If the parent doesn't act quickly and with force, the child might get the upper hand and, as a result, the parent might lose some control permanently. While I don't believe that a child should be able to get away with such deliberate misbehavior, I do believe there are other effective ways a parent can discipline his or her child without resorting to physical punishment.

You may wonder why I feel that other forms of discipline are preferable to physical punishment. What convinced me that spanking isn't necessary was that, in years of pediatric practice, I discovered there were many families in which the children were never spanked — and yet these children were cooperative, polite and kind. In some of these families the parents had not been physically punished in childhood, either. In others, the parents remembered the humiliation of being hit or spanked and were reacting to a conviction that the spankings they had received as children had had the wrong effect.

The reaction of the parents who don't spank their children because they themselves were spanked is worth considering because it raises the question of whether physical punishment does any harm. It is obvious that, when applied occasionally by loving parents, it can't do *much* harm — after all, millions of good men and women have been brought up in this way. But I think there are better ways of influencing children. When physical punishment is used frequently, especially by irritable or harsh parents, its unfavorable effects are noticeably multiplied. I believe physical punishment teaches children that might makes right and helps to turn some of them into bullies. Physical punishment leaves some sensitive children with a lasting resentment toward their parents for having humiliated them in this way. It encourages other children to feel that violence is not really bad and to think of physical force as a way of solving problems or settling disputes. As adults we know it is not an effective way of solving problems or settling disputes.

To me the most important reason for trying not to use physical punishment is that, if it is effective, it makes the child behave out of fear of the pain and out of fear of your anger. I think it's preferable for children to do the right thing because they love their parents and want to please them — not because they fear them. Then, as the children grow up, go to school, get jobs, marry and raise a family, they'll carry over this same attitude of getting along well in life by loving people, wanting to please them and cooperate with them — and receiving that love and cooperation in return.

What about other punishments parents can use, such as taking away a beloved toy for a day or so? To me, the loss of a privilege seems better than the indignity of being hit.

Isolating a child who is out of control has been used effectively in good day-care centers. Sending a child to his room for a given period of time works just as well at home, but isolation should be used in a calm, friendly spirit, as a way of helping the child to cool off.

To me, the best way of ensuring good behavior is for parents to show children love and respect — from infancy — and to set a good example. Then children look up to their parents and want to please them.

When parents shout and hit, they thwart a child's natural desire to please her parents, because the child's love and respect for them has been diminished. In the long run, that makes the parents' job of disciplining their children all the more difficult.

You may think your children would never respond to anything as mild as a good example or a polite request. If they have been used to rougher forms of discipline, I'll admit that they will seem insensitive at first to gentler methods. But they will gradually come around. I've seen the transformation take place in a day-care center, where a thick-skinned misbehaver began cooperating with a gentle teacher after he slowly learned that he could trust her to be kind to him.

One approach you could use to get the attention of a child who has learned to ignore anything but the most extreme forms of correction would be to go to her immediately when she misbehaves, put your arm around her and say quietly, "When you do that, it makes me unhappy. Please don't do it again!" If misbehavior is consistently corrected in this fashion, not only will the child learn that she can't persist in whatever it is that she's doing wrong, but, more importantly, she will come to enjoy a better relationship with you and the impulse to misbehave will diminish. Of course, it takes a good deal of patience for a parent to make the shift to this kind of gentle discipline. But the results are well worth the effort.

Although Dr. Spock cannot answer readers' letters individually, he will respond to them in his column. Please address your questions to Department DW, Redbook, 224 West 57th Street, New York, NY 10019.

Benjamin Spock: "Have You Ever Wanted to Strike Your Child?" Reprinted from *Redbook* by permission of the publisher.

for *plausibility of distractors.* As with the preceding task, *type of information* received an abstractness score of 4 (because the task requests a reason, rather than an alternative, an amount, or a thing).

A more difficult task (RP80 score value 346) directs the reader to identify and list two similarities between the new and old ways American Express handles charge card receipts (Exhibit 13-9). This piece of text received a Fry readability score of 8 (eighth grade), indicating that this should not be a difficult text for most adults. The correct response requires the reader to compare and contrast several pieces of information to determine these similarities. Because integrating information through comparison and contrast is relatively demanding, this task was coded 4 for *type of match*. When the information requested is a "similarity," as in this task, *type of information* is scored as a 4. *Plausibility of distractors* received a score of 2 (because similar information appears somewhere in the text, but not in the same paragraph as the answer).

Exhibit 13-9. Example of text for difficult prose literacy tasks

American Express' Way of Handling the Flood of Charge Card Receipts

How the new way stacks up against the old way

The New Way:

1 Image processing camera converts receipts to electronic digital image and paper receipts are discarded. **2** Digital image is scanned for account and invoice numbers by optical character (99% accuracy). In the future, computers will also read handwritten charge amounts. **3** Charge amounts are entered by computer operator from image displayed on computer screen. **4** Images are sorted electronically. **5** Bills, with images of receipts, are printed by laser and mailed to cardholders. **6** Images of receipts are stored permanently on optical discs.

The Old Way:

1 Paper receipts are microfilmed for **2** permanent storage, then **3** scanned for account and invoice number by optical character reader (82% accuracy). **4** Charge amounts are entered by computer operator from receipts. **5** A code containing all the information is printed on the receipts. **6** Paper receipts are sorted. **7** Bills are generated by mainframe computer. **8** Receipts and bills are joined and mailed.

One of the most difficult prose tasks in the 1990 survey (RP80 score of 441) required the reader to identify two differences in the new and old ways of handling the American Express receipts. The correct response requires the reader to compare and contrast several pieces of information to determine these differences. Because identifying differences through comparison and contrast is more demanding than

identifying similarities, this task was coded 7 for *type of match*. When the information requested is a "difference," as in this task, *type of information* is scored as a 5. This task was also coded as a 5 for *plausibility of distractors.*

The tasks shown above provide examples of how the process variables were assigned numeric values to capture the extent of the various information-processing demands that such prose literacy tasks place on readers.

13.4.1.3 Coding the 1985 and 1990 prose literacy tasks

The preceding section provided examples to illustrate how type of match, type of information, and plausibility of distractors were coded for selected prose literacy tasks from the 1990 literacy survey of job-seekers served by the Department of Labor. In this section, the coding rules are formulated in more general terms. A number of criteria must be taken into account when measuring the four variables associated with task difficulty on the prose scale.

13.4.1.3.1 Type of Match

Description. Type of match refers to the processes used to relate requested information to the corresponding information in a prose text, and to the process of entering a response. Four basic types of match can be distinguished: *locate*, *cycle*, *integrate*, and *generate* matches.

Locate tasks require users to match one or more features in a question to one or more features in the text (Kirsch and Mosenthal, 1992a; Mosenthal, 1998). Based on this match, the answer is located in the appropriate paragraph or sentence of a prose text.

Cycle tasks require users to perform an iterative series of locate matches and may involve the selection of several pieces of information that meet a criterion. With prose texts, cycle tasks are made difficult depending upon whether they are performed within a paragraph or between paragraphs. Cycle tasks are further made difficult depending upon whether the cycles are independent of one another or a sequence in which each answer is used to identify the next part of the locating cycle.

Integrate tasks require users to compare or contrast information that has been located in two or more different locate matches or in one or more cycle matches (Kirsch and Mosenthal, 1992; Mosenthal, 1998). In general, integrate tasks which require readers to compare information are easier than those that require readers to contrast information.

Generate tasks require readers to use prior knowledge (often representing a specific type of content knowledge) to match information in a question or directive to corresponding information in a prose text (Mosenthal and Kirsch, 1993a; Mosenthal, 1998). Moreover, generate tasks may require readers to use specialized knowledge to select from among a set of plausibly correct responses the answer which best

meets the conditions stated in a question or directive. Without the benefit of such knowledge, users often must guess or ask some expert to complete the match.

Scoring rules. The scoring of type of match is basically determined by the nature of the match, and to a lesser extent by the number of phrases or features in the request, the number of responses requested, whether prior knowledge is needed for the required inferences, and how a choice among several possible answers should be selected. The basic idea, however, is that locate matches are easier than cycle matches, cycle matches are easier than integrate matches, and integrate matches are easier than generate matches.

Exhibit 13-10. Basic scoring rules for type of match: Nature of the task

Rule	Score
When the task is to *locate* the information in the prose text or document that corresponds to the features requested.	1
When the task is to *cycle* (that is, perform an iterative series of locate matches) to find the information that corresponds to the features requested. Add 1 point if the answer is located in more than one paragraph.	2
When the task is to *integrate* information located in a prose text by comparing, or for prose text, when the task is to infer a condition based on a synthesis of features found in the same paragraph of text.	3
When the task is to integrate information located in a prose text by contrasting, or for prose text, when the task is to infer a condition based on a synthesis of features found in more than one paragraph of text.	4
When the task is to *generate* new information (that is, to use prior knowledge to match information requested with that in the prose text).	5

Sometimes matching is made more difficult as the number of *phrases or features* in the directions required to locate an answer increases (Mosenthal, 1998). Matches that require the identification of only a single phrase or feature are, on average, easier than matches that require the identification of two phrases or features. Given the basic score based on the nature of the match, additional points can be added based on the number of phrases or features in the directions.

Exhibit 13-11. Additional scoring rule for type of match: Number of phrases or features in request

Rule	Add
When the request for information consists of one independent clause and one dependent clause	1
When the request for information consists of one independent clause and two dependent clauses	2
When the request for information consists of one independent clause and three or more dependent clauses	3

Sometimes matching is also made more difficult as the number of *responses* readers must supply increases and as the specificity of this number decreases. Requests of readers to list only one answer are easier than requests to list two or three answers; requests of readers to list two or three answers are easier than requests for four answers. Requests that do not specify the number of responses explicitly are harder than those that do specify the number of multiple responses required. Given the preliminary score based on the nature of the match and the number of phrases or features, additional points can be added based on the number of responses requested.

Exhibit 13-12. Additional scoring rule for type of match: Number of responses requested

Rule	Add
When readers are requested to list two or three answers	1
When readers are requested to list four or five answers $ add 1 point if the request does not specify exactly how many	2
when readers are requested to list six or more answers $ add 1 point if the request does not specify exactly how many	3

Matching can be further made difficult to the extent that readers have to *make inferences* to match information in the question to information in the document or text (Mosenthal and Kirsch, 1993b). Questions may require either low text-based inferences (which consist of inferences to be made within the context of information provided in a text), or high text-based inferences (which consist of inferences requiring some combination of knowledge of the text and specialized prior knowledge). Low text-based inferences are easier to make than high text-based inferences. Given the preliminary score based on the nature of the match, the number of phrases or features, and the number of responses requested, additional points can be added based on the kind of inference needed to answer.

Exhibit 13-13. Additional scoring rule for type of match: Inferences needed

Rule	Add
When the text alone provides sufficient information to make an inference needed to match the request with the information in the text or document (a low text-based inference)	1
When prior knowledge as well as the text is needed to make the inference needed to match the request with the information in the text or document (a high text-based inference)	3

Sometimes matching is made more difficult when all possibilities match the request and readers have to choose which one of several possible answers best completes a requested *information frame* (Mosenthal and Kirsch, 1991). In these cases, the match is more difficult when this choice requires using the text to infer why one of several possible answers best completes a requested information frame, when this choice requires identifying conditional information which renders one of the possible answers more consistent with the conditions requested than others, or when this choice requires readers to relate a

pronoun to its antecedent before an answer can be provided. Given the preliminary score based on the nature of the match, the number of phrases or features, the number of responses requested, and the kind of inference needed, additional points can be added based on how the reader must complete an information frame.

Exhibit 13-14. Additional scoring rule for type of match: Completing an information frame

Rule	Add
When the choice among candidate answers requires a low, text-based inference (the text alone provides sufficient information to make the inference), the identification of a condition, the identification of a pronoun antecedent, or a restatement of a type of information	2
When the choice among candidate answers requires a high, text-based inference (prior knowledge as well as the text is needed to make the inference)	4

These scoring rules are additive (Meyer, Marsiske, and Willis, 1993). A prose literacy task, for example, might have a basic score of 2 because it is a cycle task, but have additional points added because the cycling occurs between paragraphs (add 1), involves a two-clause question (add 1), needs a low text-based inference (add 1), for which the answer should consist of two responses (add 1), but whose actual number is not explicitly specified (add 1). A prose assessment task with these features would have a total type-of-match score of 7.

The actual prose-based tasks used in this assessment scored from 1 to 8 on type of match. Eight was not a ceiling set in advance. Rather, these upper bounds reflects the range of difficulty combinations which commonly characterize tasks found in society and the workplace. While more difficult tasks could be conceived in designing assessments (for example, a four-phrase contrast task requiring high text-based inferencing and six uncued responses), such tasks would be so difficult that they would bear little resemblance to ordinary usage of prose texts.

13.4.1.3.2 Type of Information

Description. Type of information refers to the degree of concreteness of the objects described in a prose text or document (Mosenthal and Kirsch, 1991b; Mosenthal, 1998). More abstract objects are harder for readers to identify and understand.

Scoring rules. Assessment tasks can be rated in terms of the concreteness of the information requested. Most concrete were tasks requesting information about persons, groups, animals, locations, and things. Somewhat less concrete were questions requiring the identification of amounts, times, attributes, types, actions, locations, and parts. Somewhat abstract were questions requesting information about manner, goals, purposes, alternatives, conditions, pronoun references, and predicate adjectives. Abstract

tasks requested the identification of causes, effects, reasons, evidence, similarities, and explanations. Finally, very abstract tasks requested the identification of equivalences, differences, themes, or patterns.

Exhibit 13-15. Scoring rules for type of information

Rule	Score
When the information requested refers to a person, group, animal, place, or thing (most concrete)	1
When the information requested refers to an amount, time, attribute, type, action, location, or part	2
When the information requested refers to a manner, goal, purpose, alternative, attempt, condition, pronominal reference, predicate adjective, sequence, assertion, problem, solution, role, or process	3
When the information requested refers to a cause, effect, reason, result, evidence, similarity, explanation, opinion, or procedure	4
When the information requested refers to an equivalence, difference, theme, pattern, definition, or advantage (most abstract)	5

13.4.1.3.3 Plausibility of Distractors

Description. Distractors are elements of a question's given or requested information that appear in the prose text that, when identified lack an essential feature and do not qualify as correct. Given information is provided by the directions for a task and is used to search for the requested information (that being sought). Unless the possible but incorrect answers are plausible (by sharing some of the features of the correct answer), they do not function as distractors (Mosenthal, 1996, 1998).

Scoring rules. Tasks are easiest when the prose text contains no information related to the conditions set forth in the question other than the answer. Tasks become slightly more difficult when a distractor for either given or requested information (but not both) appears, but does not occur very close to the correct answer. Tasks become more difficult when plausible distractors for both given and requested information appear, but are not both located next to the correct information. This occurs in prose texts when they appear in different paragraphs, one of which may be in the paragraph in which the answer occurs.

Tasks become still more difficult when plausible distractors for both given and requested information appear in the same place but are not located near the correct information. This occurs in prose texts when they both appear in the same paragraph but one other than the paragraph in which the answer appears. Tasks involving prose texts also reach this level of difficulty when negatives become involved—the plausible distractors represent the opposite condition of what is established in the question or directive, and these distractors appear in a paragraph other than the one in which the answer occurs.

Tasks are most difficult when plausible distractors for both given and requested information appear in the same place, or appear with negative conditions, and are located near the correct information. This occurs in prose texts when they both appear in the same paragraph as the one in which the answer occurs, or when the distractors represent the opposite condition of what is established by the task, and they appear in the same paragraph as the answer.

Exhibit 13-16. Scoring rules for plausibility of distractors

Rules for prose texts	Score
When no information related to the conditions requested appears, other than the answer (no plausible distractors)	1
When information similar to either given or requested information appears somewhere in the text but not near the answer, or inferences invited by information in the paragraph containing the answer bear a resemblance to the answer	2
When distractors for both given and requested information appear in different paragraphs, though one could occur in the paragraph containing the answer	3
When distractors for both given and requested information, or when plausible distractors represent the opposite condition of what is requested, appear in the same paragraph, but one other than the paragraph containing the answer	4
When distractors for both given and requested information, or when plausible distractors represent the opposite condition of what is requested, appear in the same paragraph as the answer	5

13.4.1.3.4 Readability

Description. Prose texts vary in the length of sentences, number of syllables in the words used, and the complexity of the syntax, while documents vary in their complexity, depending on their organization, number of elements, and number of labels. Literacy tasks may be easier to process when the structure of the document or prose text containing the needed information is less complex. The measurement of the complexity of prose texts derives from Fry's research on readability (Fry, 1975, 1977, 1981).

Scoring rules for prose text. Readability of prose is based on the average number of syllables per 100 words and the average number of sentences per 100 words. These two continuous variables are then used as coordinates in Fry's (1977) readability grade level graph, which portrays a nonlinear relationship between the two and the resulting readability level. In general, however, the more syllables per word and the more words per sentence, the higher the associated grade level of the text. The grade levels of the texts used in the National Adult Literacy Survey ranged from fourth to fifteenth.

13.4.1.4 Codes for all 1985 and 1990 prose literacy tasks

The preceding sections described in detail the several criteria that must be taken into account when measuring the four variables associated with task difficulty on the prose scale. These rules were applied to

all prose literacy tasks in the 1990 survey and in the 1985 young adult literacy assessment. The resulting codes, along with RP80 task difficulties and IRT item parameters are shown in Table 13-2.

Table 13-2. List of prose literacy tasks, along with RP80 task difficulty, IRT item parameters, and values of variables associated with task difficulty: 1990 survey of the literacy of job-seekers

Identifier	Task Description	Scaled RP80	IRT parameters			Read-ability	Type of match	Distractor Plausibility	Information type
			a	b	c				
A111301	Toyota, Acura, Nissan	189	0.868	-2.488	0.000	8	1	1	1
AB21101	Swimmer: Underline sentence telling what Ms. Chanin ate	208	1.125	-1.901	0.000	8	1	1	1
A120501	Blood donor pamphlet	216	0.945	-1.896	0.000	7	1	1	2
A130601	Summons for jury service	237	1.213	-1.295	0.000	7	3	2	2
A120301	Blood donor pamphlet	245	0.956	-1.322	0.000	7	1	2	3
A100201	PHP subscriber letter	249	1.005	-1.195	0.000	10	3	1	3
A111401	Toyota, Acura, Nissan	250	1.144	-1.088	0.000	8	3	2	4
A121401	Dr. Spock column: alterntv to phys punish	251	1.035	-1.146	0.000	8	2	2	3
AB21201	Swimmer: Age Ms. Chanin began to swim competitively	250	1.070	-1.125	0.000	8	3	4	2
A131001	Shadows Columbus saw	280	1.578	-0.312	0.000	9	3	1	2
AB80801	Illegal questions	265	1.141	-0.788	0.000	6	3	2	2
AB41001	Declaration: Describe what poem is about	263	0.622	-1.433	0.000	4	3	1	3
AB81101	New methods for capital gains	277	1.025	-0.638	0.000	7	4	1	3
AB71001	Instruction to return appliance: Indicate best note	275	1.378	-0.306	0.266	5	3	2	3
AB90501	Questions for new jurors	281	1.118	-0.493	0.000	6	4	2	1
AB90701	Financial security tips	262	1.563	-0.667	0.000	8	3	2	4
A130901	Shadows Columbus saw	282	1.633	-0.255	0.000	9	3	4	1
AB60201	Make out check: Write letter explaining bill error	280	1.241	-0.440	0.000	7	3	2	4
AB90601	Financial security tips	299	1.295	-0.050	0.000	8	2	2	4
A121201	Dr. Spock column: why phys punish accptd	285	1.167	-0.390	0.000	8	3	2	4
AB70401	Almanac vitamins: List correct info from almanac	289	0.706	-0.765	0.000	7	3	4	1
A100301	PHP subscriber letter	294	0.853	-0.479	0.000	10	4	3	2
A130701	Shadows Columbus saw	298	1.070	-0.203	0.000	9	3	2	3
A130801	Shadows Columbus saw	303	0.515	-0.929	0.000	9	3	2	2
AB60601	Economic index: Underline sent. Explaining action	305	0.809	-0.320	0.000	10	3	2	4
A121301	Dr. Spock column: 2 cons against phys punish	312	0.836	-0.139	0.000	8	3	3	4
AB90401	Questions for new jurors	300	1.230	-0.072	0.000	6	4	2	3
AB80901	Illegal questions	316	0.905	0.003	0.000	6	4	3	3
A111101	Toyota, Acura, Nissan	319	0.772	-0.084	0.000	8	4	3	2
AB40901	Korean Jet: Give argument made in article	329	0.826	0.166	0.000	10	4	4	4
A131101	Shadows Columbus saw	332	0.849	0.258	0.000	9	5	4	1
AB90801	Financial security tips	331	0.851	0.236	0.000	8	5	5	2
AB30601	Technology: Orally explain info from article	333	0.915	0.347	0.000	8	4	4	4
AB50201	Panel: Determine surprising future headline	343	1.161	0.861	0.196	13	4	4	4
A101101	AmerExp: 2 Similarities in Handling Receipts	346	0.763	0.416	0.000	8	4	2	4
AB71101	Explain difference between 2 types of benefits	348	0.783	0.482	0.000	9	6	2	5
AB81301	New methods for capital gains	355	0.803	0.652	0.000	7	5	5	3
A120401	Blood donor pamphlet	358	0.458	-0.056	0.000	7	4	5	2
AB31201	Dickinson: Describe what is expessed in poem	363	0.725	0.691	0.000	6	6	2	4
AB30501	Technology: Underline sentence explaining action	371	0.591	0.593	0.000	8	6	4	4
AB81201	New methods for capital gains	384	0.295	-0.546	0.000	7	2	4	2
A111201	Toyota, Acura, Nissan	404	0.578	1.192	0.000	8	8	4	5
A101201	AmExp: 2 Diffs in Handling Receipts	441	0.630	2.034	0.000	8	7	5	5
AB50101	Panel: Find information from article	469	0.466	2.112	0.000	13	6	5	4

13.4.1.5 Validity evidence for the prose scale

One important piece of validation evidence can be obtained from analyses of the tasks used in the 1990 job-seeker survey. As indicated earlier, the prose area was not especially well defined in the 1985 young adult literacy assessment—the scale contained only 15 tasks. As a result, about 30 new prose tasks were developed for the 1990 survey. Each of the new tasks was coded using the revised theoretical framework described in the preceding section. These variables were then used in regression analyses designed to predict the scale values of each task. Table 13-3 shows the results of these analyses for both the new tasks as well as the entire 1990 pool of prose literacy tasks (1985 and 1990 tasks). The numbers in the tables represent the raw beta coefficients for each of the variables included in the regression analyses along with standard errors and probabilities. Overall, the three process variables were significant for both the new 1990 tasks and for the complete set of prose tasks. Although not shown here, readability was significant if entered into the regression by itself and accounted for about 20 percent of the variance in predicting task scale values. However, when combined with the three process variables, it did not increase the explained variance.

Table 13-3. Unstandardized regression coefficients and standard errors predicting RP80 task difficulties on the basis of four structure and process variables: 30 new prose literacy tasks and 44 total prose literacy tasks from the 1990 survey of job-seekers

	New Tasks			All Tasks		
	Coeff	StdErr	p	Coeff	StdErr	p
Structure Variable						
Readability	-.04	3.97	.93	2.14	2.35	.37
Process Variables						
Type of match	12.08	5.94	.05	17.68	4.48	.00
Plausibility of distractors	28.53	5.79	.00	20.98	4.56	.00
Type of information	14.35	5.16	.01	12.19	4.28	.01
Variance accounted for:						
R^2		81%			82%	
Adjusted R^2		78%			80%	
Degrees of freedom		25			39	

Using the expanded theoretical framework for task development on the prose scale appears to have been successful**.** The amount of variance accounted for in the new tasks (81 percent), as well as in the complete set of tasks (82 percent), compares favorably with the research results reported by Kirsch and Mosenthal (1990)—in the range from 81 to 89 percent for document tasks. These results also suggest that readability is less important than the process variables in explaining task difficulty.

13.4.2 Document Literacy

An additional aspect of being literate in today's society is having the knowledge and skills needed to process documents, or information organized in matrix structures (i.e., in rows and columns). Included among documents are such things as tables, signs, indexes, lists, coupons, schedules, charts, graphs, maps, and forms. In contrast to prose, which tends to be the predominant form of literacy in schools, documents tend to be the principal form of literacy in non-school settings (Guthrie, Seifert, and Kirsch, 1986). Documents serve many important functions in our daily lives (Bassett, Goodman, and Fosegan, 1981; Burch and Grudnitski, 1986). They enable people to perform important actions (e.g., applying for benefits, opening a charge account), make informed decisions (e.g., using a table of benefits to determine whether certain medical costs are covered), and record actions (e.g., completing a deposit slip or bill of sale, receiving a ticket for speeding).

13.4.2.1 Document variables

Document literacy tasks require readers to locate and match information in a question or directive to corresponding information in complex arrays, and to use this information in appropriate ways. For example, procedural knowledge may be needed to transfer information from one source or document to another, as is necessary in completing applications or order forms. This matching again involves the strategies of *locating*, *cycling*, *integrating*, and *generating* information; these strategies are again represented by the variable *type of match* (Mosenthal, 1996). As with prose literacy tasks, success in processing documents also appears to depend on the ability to identify different *types of information*. Similarly, both prose and document tasks are made more difficult through the presence of *plausible distractors*.

In addition to these three process variables, it was deemed important to provide an index of the readability of document stimuli. Since no such index is readily available for estimating document complexity, Mosenthal and Kirsch (1998) developed a means of estimating this complexity based on the grammar used in the earlier research.

The basic structural unit of documents is "simple lists" (Mosenthal and Kirsch, 1989a, 1998). Such lists consist of a series of exemplars or items that belong to a common class of elements that, in most instances, are organized in terms of a more generic category or label. The documents used in the assessments reflect the ways in which a number of simple lists are organized to present more interdependent sets of information. These lists have been described as combined, intersecting, and nested (Kirsch and Mosenthal, 1989, 1990b; Mosenthal and Kirsch, 1989b, 1998). The document readability variable ranges from 1 to 11. Included in this score is a number reflecting the type of document structure, the number of labels, and the numbers of items.

13.4.2.2 Examples of document literacy tasks

One of the easier tasks on the document scale (RP80 score of 198) required the reader to look at a theater trip notice containing information about two plays (Exhibit 13-17). This document received a low structural complexity score of 2. The reader is directed to circle the cost for a ticket and bus trip to see *On the Town*. Although the reader simply locates the line in the notice labeled "price" and circles the dollar amount associated with *On the Town*, the cost given in the document for *Sleuth* serves as a plausible distractor. This task received a code of 1 for *type of match*, and codes of 2 for *plausibility of distractors* and *type of information*.

Exhibit 13-17. Example of stimulus material for a relatively easy document literacy task

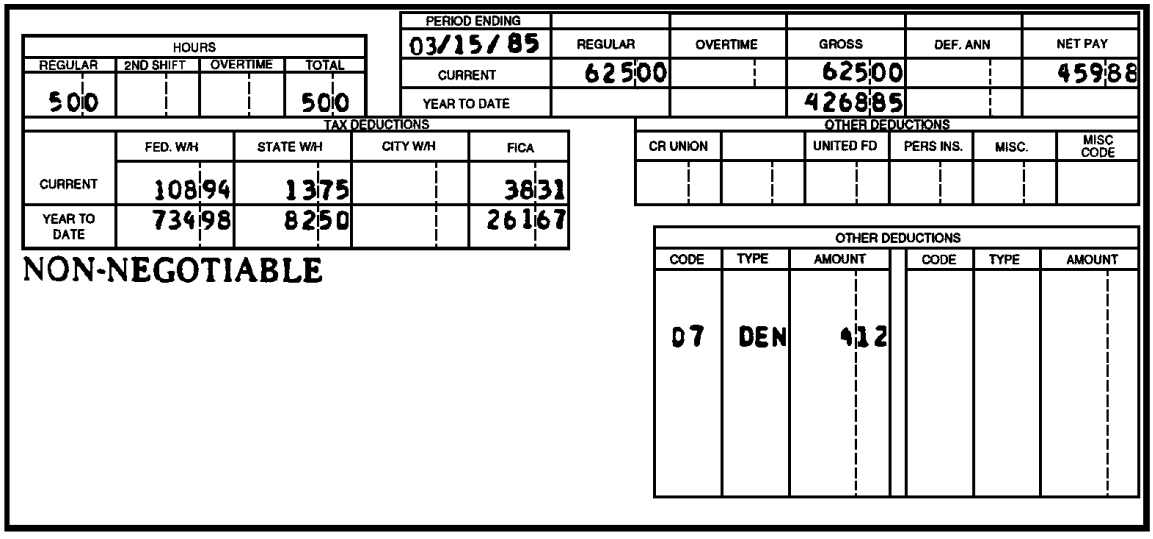
THEATER TRIP

A charter bus will leave from the bus stop (near the Conference Center) at 4 p.m., giving you plenty of time for dinner in New York. Return trip will start from West 45th street directly following the plays. Both theaters are on West 45th Street. Allow about 1 1/2 hours for the return trip.

Time:	**4 p.m., Saturday, November 20**		
Price:	**"On the Town"**	**Ticket and bus**	**$11**
	"Sleuth"	**Ticket and bus**	**$8.50**
Limit:	**Two tickets per person**		

A more difficult task at 275 on the document scale directs the reader to look at a wage and tax statement (Exhibit) and to select "gross pay for this year to date." If readers fail to identify and match on both features—gross pay and year to date—they are likely to respond with an incorrect amount based on distracting information. The structural complexity of this document was coded 5. It was rated 2 on *type of match* and *type of information* with a 3 for *plausibility of distractors.*

Exhibit 13-18. Example of stimulus material for moderately difficult document literacy tasks



Another question using this wage and tax statement was also expected to require a two-feature match—current and net pay—and, therefore, to have approximately the same scale value. However, this task was considerably easier (224) and the codes assigned to the process variables indicate that little distracting information was present in the document. Each variable received a code value of 2.

Another task of similar difficulty (234) directs the reader to look at a pediatric dosage chart (Exhibit 13-19) and underline the sentence that indicates how often the medication may be administered. To respond successfully, the reader needs to associate the word "administered" in the directive to the word "given" in the document by looking at information outside the table itself. The structural complexity of this document was coded a 5. While *type of match* was coded 2, both *plausibility of distractors* and *type of information* received codes of 3.

Exhibit 13-19. Example of stimulus for moderately difficult document tasks

Recommend

ALCOHOL-FREE
ASPIRIN-FREE
Tempra®
ACETAMINOPHEN

A Caring Sponsor of
Ronald McDonald House®
Ronald McDonald House is a program of
Ronald McDonald Children's Charities

Pediatric Dosage Chart Drops, Syrup, & Chewables

Age	Approximate Weight Range*	Dosage			
		Drops	Syrup	Chewables 80 mg	Chewables 160 mg
† Under 3 mo	Under 13 lb	½ dropper	¼ tsp	—	—
† 3 to 9 mo	13-20 lb	1 dropper	½ tsp	—	—
† 10 to 24 mo	21-26 lb	1 ½ droppers	¾ tsp	—	—
2 to 3 yr	27-35 lb	2 droppers	1 tsp	2 tablets	—
4 to 5 yr	36-43 lb	3 droppers	1 ½ tsp	3 tablets	1 ½ tablets
6 to 8 yr	44-62 lb	—	2 tsp	4 tablets	2 tablets
9 to 10 yr	63-79 lb	—	2 ½ tsp	5 tablets	2 ½ tablets
11 yr	80-89 lb	—	3 tsp	6 tablets	3 tablets
12 yr and older	90 lb & over	—	3-4 tsp	6-8 tablets	3-4 tablets

† Consult with physician before administering to children under the age of 2 years.

Dosage may be given every 4 hours as needed but not more than 5 times daily.

How Supplied:

Drops: Each 0.8 ml dropper contains 80 mg (1.23 grains) acetaminophen.

Syrup: Each 5 ml teaspoon contains 160 mg (2.46 grains) acetaminophen.

Chewables: Regular tablets contain 80 mg (1.23 grains) acetaminophen each. Double strength tablets contain 160 mg (2.46 grains) acetaminophen each.

* If child is significantly under- or overweight, dosage may need to be adjusted accordingly.

The weight categories in this chart are designed to approximate effective dose ranges of 10-15 milligrams per kilogram.
(Current Pediatric Diagnosis and Treatment. 8th ed. CH Kempe and HK Silver, ed. Lange Medical Publications; 1984, p. 1079.)

LA-1451-2-88 © 1988, Bristol-Myers U.S. Pharmaceutical and Nutritional Group • Evansville, Indiana 47721 U.S.A.

(c) 1988, Bristol-Myers Pharmaceutical and Nutritional Group.

A second and more difficult question (327) using the medicine dosage chart directs the reader to determine from the chart how much syrup is recommended for a child who is 10 years old and weighs 50 pounds. This task is difficult because one can not simply match literal or synonymous information to perform successfully since the weight as given in the question is less than that of the typical 10 year old according to the table. Instead, one must rely on prior knowledge, or to find the asterisked note relating to the column headed "Approximate Weight Range," that the correct dosage is to be based on weight not age to ensure that the child receives an effective dose. In any event, if the reader approaches this task as a single literal match, the age of the child is a highly plausible distractor and may lead to an incorrect response. The variable codes reflect this line of reasoning with ratings of 4 for *type of match* and *plausibility of distractors* while *type of information* is rated 2.

13.4.2.3 Coding the 1985 and 1990 document literacy tasks

The preceding section provided examples to illustrate how type of match, type of information, and plausibility of distractors were coded for selected document literacy tasks from the 1990 literacy survey of job-seekers served by the Department of Labor. In this section, the coding rules are formulated in more general terms. A number of criteria must be taken into account when measuring the four variables associated with task difficulty on the document literacy scale.

13.4.2.3.1 Type of Match

Description. Type of match refers to the processes used to relate requested information to the corresponding information in a document, and to the process of entering a response. As with the prose literacy scale, locate, cycle, integrate, and generate tasks can be distinguished. Locate tasks on the document literacy scale require users to match one or more features in a question to one or more features in the document (Kirsch and Mosenthal, 1992a; Mosenthal, 1996). Based on this match, the answer is located in the appropriate node of a document. In a document, a node is either an element of a list, a matrix cell defined by the intersection of two or more lists, or a list itself. Cycle tasks, like those on the prose literacy scale, require users to perform an iterative series of locate matches and may involve the selection of several pieces of information that meet a criterion. With documents, cycle tasks are made difficult depending upon whether they are performed within a given list or between lists. Integrate tasks require users to compare or contrast information that has been located in two or more different locate matches or in one or more cycle matches (Kirsch and Mosenthal, 1992). Generate tasks require readers to use prior knowledge (often representing a specific type of content knowledge) to match information in a question or directive to corresponding information in a document (Mosenthal and Kirsch, 1993a, 1998).

Scoring rules. The scoring of type of match is basically determined by the nature of the match, and to a lesser extent by the number of phrases or features in the request, the number of responses requested,

whether prior knowledge is needed for the required inferences, and how a choice among one of several possible answers should be selected.

Exhibit 13-20. Basic scoring rules for type of match: Nature of the task Rule	Score
When the task is to *locate* the information in the document that corresponds to the features requested.	1
When the task is to *cycle* (that is, perform an iterative series of locate matches) to find the information that corresponds to the features requested. Add 1 point if the answer identified in one match is used to carry out a second match.	2
When the task is to *integrate* information located in a document by comparing	3
When the task is to integrate information located in a document by contrasting	4
When the task is to *generate* new information (that is, to use prior knowledge to match information requested with that in the document).	5

Sometimes matching is made more difficult as the number of *phrases or features* in the directions required to locate an answer increases (Kirsch and Mosenthal, 1990; Mosenthal, 1996). Matches that require the identification of only a single phrase or feature are, on average, easier than matches that require the identification of two phrases or features.

Exhibit 13-21. Additional scoring rule for type of match: Number of phrases or features in request Rule	Add
When the request for information consists of two features	1
When the request for information consists of three features	2
When the request for information consists of four or more features	3

Sometimes matching is also made more difficult as the number of *responses* readers must supply increases and as the specificity of this number decreases. Requests of readers to list only one answer are easier than requests to list two or three answers; requests of readers to list two or three answers are easier than requests for four answers. Requests that do not specify the number of responses explicitly are harder than those that specify the number of multiple responses required.

Exhibit 13-22. Additional scoring rule for type of match: Number of responses requested Rule	Add
When readers are requested to list two or three answers	1
When readers are requested to list four or five answers $ add 1 point if the request does not specify exactly how many	2
When readers are requested to list six or more answers $ add 1 point if the request does not specify exactly how many	3

Matching can be made even more difficult to the extent that readers have to *make inferences* to match information in the question to information in the document (Mosenthal and Kirsch, 1993b, 1998). Questions may require either low text-based inferences, which consist of inferences which can be made within the context of information provided in a text, or high text-based inferences, which consist of inferences which require some combination of knowledge of the text and specialized prior knowledge.

Exhibit 13-23. Additional scoring rule for type of match: Inferences needed

Rule	Add
When the text alone provides sufficient information to make an inference needed to match the request with the information in the text or document (a low text-based inference)	1
When prior knowledge as well as the text is needed to make the inference needed to match the request with the information in the text or document (a high text-based inference)	3

Sometimes matching is made more difficult when all possibilities match the request and readers have to choose which one of several possible answers best completes a requested *information frame* (Mosenthal and Kirsch, 1991, 1996). In these cases, the match is more difficult when this choice requires using the text to infer why one of several possible answers best completes a requested information frame, when this choice requires identifying conditional information which renders one of the possible answers more consistent with the conditions requested than others, or when this choice requires readers to relate a pronoun to its antecedent before an answer can be provided.

Exhibit 13-24. Additional scoring rule for type of match: Completing an information frame

Rule	Add
When the choice among candidate answers requires a low, text-based inference (the text alone provides sufficient information to make the inference), the identification of a condition, the identification of a pronoun antecedent, or a restatement of a type of information	2
When the choice among candidate answers requires a high, text-based inference (prior knowledge as well as the text is needed to make the inference)	4

These scoring rules are additive (Meyer, Marsiske, and Willis, 1993). A document literacy task, for another example, might have a basic score of 1 because it is a locate task, but have additional points added because two simultaneous features must be matched (add 1), and the answer should consist of three responses (add another 1) whose actual number is also not explicitly specified (add 1). A document assessment task with these features would have a total type-of-match score of 4.

While the scoring system for type of match could theoretically generate scores as high as 20, this was not the case with the 1985 and 1990 document literacy tasks. The actual document-based tasks used

scores from 1 to 8 on type of match. Tasks beyond this level, while possible, would be so difficult that they would bear little resemblance to the ordinary usage of documents.

13.4.2.3.2 Plausibility of Distractors

Description. Distractors are elements of a question's given or requested information that appear in the prose text or document, but when identified lack an essential feature and do not qualify as correct. Unless the possible but incorrect answers share some of the features of the correct answer, they do not function as distractors. Defining how close the distractor is to the correct answer involves a series of decision rules that differ somewhat depending on whether the material occurs in the form of prose texts or documents (Mosenthal, 1996, 1998).

Scoring rules. Tasks are easiest when the document contains no information related to the conditions set forth in the question other than the answer. Tasks are also easiest when there is only a single item in a list, or there is only one list with a unique label unrelated to the other labels in a document.

Tasks become slightly more difficult when a distractor appears, but does not occur very close to the correct answer. This occurs in documents when there is more than one item in a list in which one is searching for requested information or when there are labels in other lists that bear a resemblance to the label on which one is searching.

Tasks become more difficult when plausible distractors for both given and requested information appear, but are not both located next to the correct information. This occurs in documents when they appear in different matrix cells or in lists other than the cell or list in which an answer actually appears.

Tasks become still more difficult when plausible distractors for both given and requested information appear in the same place but are not located near the correct information. This occurs in documents when one or more features from both appear in a matrix cell or list other than the one in which the answer appears.

Tasks are most difficult when plausible distractors for both given and requested information appear in the same place, or appear with negative conditions, and are located near the correct information. This occurs in documents when one or more features from both requested and given information appear in the same matrix cell or list as the answer.

Exhibit 13-25. Scoring rules for plausibility of distractors	
Rule	Score
When no information related to the conditions requested appears, other than the answer (no plausible distractors)	1
When other information somewhere in a document (but not near the answer) bears a resemblance to the requested information	2
When distractors for both given and requested information appear in different matrix cells or in lists other than the cell or list containing an answer	3
When distractors for both given and requested information appear in a matrix cell or list other than the answer node	4
When distractors for both given and requested information appear in the same matrix cell or list as the answer	5

13.4.2.3.3 Type of Information

Description. Type of information refers to the degree of concreteness of the objects described in a prose text or document (Mosenthal and Kirsch, 1991b). More abstract objects are harder for readers to identify and understand.

Scoring rules. The scoring rules for type of information were identical for prose texts and documents, so the same procedures were used.

Exhibit 13-26. Scoring rules for type of information	
Rule	Score
When the information requested refers to a person, animal, place, or thing (most concrete)	1
When the information requested refers to an amount, time, attribute, type, action, or location	2
When the information requested refers to a manner, goal, purpose, alternative, attempt, condition, pronominal reference, or predicate adjective	3
When the information requested refers to a cause, effect, reason, result, evidence, similarity, or explanation	4
When the information requested refers to an equivalence, difference, theme, or pattern (most abstract)	5

13.4.2.3.4 Structural Complexity

Description. Just as prose texts vary in the length of sentences, number of syllables in the words used, and the complexity of the syntax, documents also vary in their complexity. The complexity of documents depends on their organization, number of elements, and number of labels. The measurement of the complexity of documents derives from research by Mosenthal and Kirsch (1989, 1991a, 1998) and their predecessors (Meyer and Rice, 1984).

Scoring rules for documents. Structural complexity of documents is based on the type of document, along with the number of items and labels comprising the document. The basic structural unit of documents is the simple list, which consists of a series of items or elements that belong to a common, more generic category called a "label." More complex documents are combinations of simple lists: combined, intersected, nested, or multiple lists, depending on their relationships to one another. After the initial score is derived from the document type, a supplementary set of rules is applied concerning the number of items and labels comprising the document.

Exhibit 13-27. Basic scoring rules for structural complexity: Type of document

Rule	Score
When the document has a simple list structure.	1
When the document has a combined list structure.	2
When the document has an intersected list structure.	3
When the document has a nested list structure.	4
When the document consists of different multiple documents or a combined list with minimal structure.	5

Next, the items in each list are counted. In most cases, items are the cells or basic elements of any given list, but for lists that consist of a series of sentences, each independent and dependent clause within each sentence is counted as a separate item. As shown in the table below, the additional credit to be added to the basic score depends on the number of items.

Exhibit 13-28. Additional scoring rule for structural complexity: Number of items in lists

Rule	Add
When the lists consist of 76-125 items	1
When the lists consist of 126-175 items	2
When the lists consist of more than 175 items	3

Finally, the labels heading each list are counted. Labels are the list headings that describe the contents of the list. Complex documents containing more than one list can have many labels for different parts of the document. As shown in the table below, the additional credit to be added to the basic score depends on the number of labels.

Exhibit 13-29. Additional scoring rule for structural complexity: Number of labels in lists	
Rule	Add
When the document contains 16-25 labels	1
When the document contains 26-35 labels	2
When the document contains more than 35 labels	3

These scoring rules are additive. For example, a document might be assigned a structural complexity score of 2 because it is a combined list (begin with 2 points) which involves 100 items (add 1) in 3 labeled columns (don't add anything). The total structural complexity score, given the combined features of the document, is 3.

13.4.2.4 Codes for all 1985 and 1990 document literacy tasks

The preceding sections described in detail the several criteria that must be taken into account when measuring the four variables associated with task difficulty on the document literacy scale. These rules were applied to all document literacy tasks in the 1990 survey and in the 1985 young adult literacy assessment, and the resulting codes, along with RP80 task difficulties and IRT item parameters are shown in Table 13-4.

Table 13-4. List of document literacy tasks, along with RP80 task difficulty score, IRT item parameters, and values of variables associated with task difficulty (structural complexity, type of match, plausibility of distractor, type of information): 1990 survey of the literacy of job-seekers

Identifier	Task Description	RP80	IRT parameters			Complexity	Match	Distractor	Information
			a	b	c				
SCOR100	Social Security card: Sign name on line	70	0.505	-4.804	0.000	1	1	1	1
SCOR300	Driver's license: Locate expiration date	152	0.918	-2.525	0.000	2	1	2	1
SCOR200	Traffic Signs	176	0.566	-2.567	0.000	1	1	1	1
AB60803	Nurses' convention: What is time of program?	181	1.439	-1.650	0.000	1	1	1	1
AB60802	Nurses' convention: What is date of program?	187	1.232	-1.620	0.000	1	1	1	1
SCOR400	Medicine dosage	186	0.442	-2.779	0.000	2	1	2	2
AB71201	Mark correct movie from given information	189	0.940	-1.802	0.000	8	2	2	1
A110501	Registration & tuition Info	189	0.763	-1.960	0.000	3	1	2	2
AB70104	Job application: Complete personal information	193	0.543	-2.337	0.000	1	2	1	2
AB60801	Nurses' convention: Write correct day of program	199	1.017	-1.539	0.000	1	1	2	1
SCOR500	Theatre trip information	197	0.671	-1.952	0.000	2	1	2	2
AB60301	Phone message: Write correct name of caller	200	1.454	-1.283	0.000	1	1	2	1
AB60302	Phone message: Write correct number of caller	202	1.069	-1.434	0.000	1	1	1	1
AB80301	How companies share market	203	1.292	-1.250	0.000	7	2	2	2
AB60401	Food coupons	204	0.633	-1.898	0.000	3	2	2	1
AB60701	Nurses' convention: Who would be asked questions	206	1.179	-1.296	0.000	1	2	2	1
A120601	MasterCard/Visa statement	211	0.997	-1.296	0.000	6	1	2	2
AB61001	Nurses' convention: Write correct place for tables	217	0.766	-1.454	0.000	1	1	2	2
A110301	Dessert recipes	216	1.029	-1.173	0.000	5	3	2	1
AB70903	Checking deposit: Enter correct amount of check	223	1.266	-0.922	0.000	3	2	2	1
AB70901	Checking deposit: Enter correct date	224	0.990	-1.089	0.000	3	1	1	1
AB50801	Wage & tax statement: What is current net pay?	224	0.734	-1.366	0.000	5	2	2	2
A130201	El Paso Gas & Electric bill	223	1.317	-0.868	0.000	8	1	2	2
AB70801	Classified: Match list with coupons	229	1.143	-0.881	0.000	8	2	3	1
AB30101	Street map: Locate intersection	232	0.954	-0.956	0.000	4	2	2	2
AB30201	Sign out sheet: Respond to call about resident	232	0.615	-1.408	0.000	2	3	2	1
AB40101	School registration: Mark correct age information	234	0.821	-1.063	0.000	6	2	2	3
A131201	Tempra dosage chart	233	1.005	-0.872	0.000	5	2	3	3
AB31301	Facts about fire: Mark information in article	235	0.721	-1.170	0.000	1	2	3	2
AB80401	How companies share market	236	1.014	-0.815	0.000	7	3	2	2
AB60306	Phone message: Write whom message is for	237	0.948	-0.868	0.000	1	2	3	1
AB60104	Make out check: Enter correct amount written out	238	1.538	-0.525	0.000	6	3	2	1
AB21301	Bus schedule	238	0.593	-1.345	0.000	2	2	3	2
A110201	Dessert recipes	239	0.821	-0.947	0.000	5	3	2	1
AB30301	Sign out sheet: Respond to call about resident	240	0.904	-0.845	0.000	2	2	2	3
AB30701	Major medical:locate Eligibility from table	245	0.961	-0.703	0.000	4	2	2	2
AB60103	Make out check: Enter correct amount in numbers	245	0.993	-0.674	0.000	6	3	2	1
AB60101	Make out check: Enter correct date on check	246	1.254	-0.497	0.000	6	3	2	1
AB60102	Make out check: Paid to the correct place	246	1.408	-0.425	0.000	6	3	2	1
AB50401	Catalog order: Order product one	247	0.773	-0.883	0.000	8	3	2	1
AB60303	Phone message: Mark "please call" box	249	0.904	-0.680	0.000	1	2	2	2
AB50701	Almanac football: Explain why an award is given	254	1.182	-0.373	0.000	6	2	2	3
AB20101	Energy graph: Find answer for given conditions (1)	255	1.154	-0.193	0.228	4	3	2	1
A120901	MasterCard/Visa statement	257	0.610	-0.974	0.000	6	1	2	2
A130101	El Paso Gas & Electric bill	257	0.953	-0.483	0.000	8	2	2	2
AB91101	Minimum wage power	260	0.921	-0.447	0.000	4	3	3	2
AB81001	Consumer Reports books	261	1.093	-0.304	0.000	4	3	2	1
AB90101	Pest control warning	261	0.889	-0.471	0.000	2	3	3	2
AB21501	With graph, predict sales for spring 1985	261	0.799	-0.572	0.000	5	3	2	2
AB20601	Yellow pages: Find place open Saturday	266	1.078	-0.143	0.106	7	3	2	1
A130401	El Paso Gas & Electric bill	270	0.635	-0.663	0.000	8	3	3	2
AB70902	Checking deposit: Enter correct cash amount	271	0.858	-0.303	0.000	3	3	3	2

Table 13-4. List of document literacy tasks, along with RP80 task difficulty score, IRT item parameters, and values of variables associated with task difficulty (structural complexity, type of match, plausibility of distractor, type of information): 1990 survey of the literacy of job-seekers — Continued

Identifier	Task Description	RP80	IRT parameters			Complexity	Match	Distractor	Information
			a	b	c				
AB50601	Almanac football: Locate page of info in almanac	276	1.001	-0.083	0.000	5	3	2	2
A110701	Registration & tuition Info	277	0.820	-0.246	0.000	3	2	5	2
AB20201	Energy graph: Find answer for given conditions (2)	278	0.936	-0.023	0.097	4	4	2	1
AB31101	Abrasive gd: Can product be used in given case?	280	0.762	-0.257	0.000	10	5	2	3
AB80101	Burning out of control	281	0.550	-0.656	0.000	2	3	2	2
AB70701	Follow directions on map: Give correct location	284	0.799	-0.126	0.000	4	4	2	2
A110801	Washington/Boston schedule	284	0.491	-0.766	0.000	9	2	4	2
AB70301	Almanac vitamins: Locate list of info in almanac	287	0.754	-0.134	0.000	5	3	4	2
AB20401	Yellow pages: Find a list of stores	289	0.479	-0.468	0.144	7	2	5	1
AB20501	Yellow pages: Find phone number of given place	291	0.415	-0.772	0.088	7	2	4	2
AB60305	Phone message: Write who took the message	293	0.640	-0.221	0.000	1	5	2	1
AB30401	Sign out sheet: Respond to call about resident (2)	297	0.666	-0.089	0.000	2	2	1	4
AB31001	Abrasive guide: Type of sandpaper for sealing	304	0.831	0.285	0.000	10	4	2	2
AB20301	Energy: Yr 2000 source prcnt power larger than 71	307	1.090	0.684	0.142	4	4	2	1
AB90901	U.S. Savings Bonds	308	0.932	0.479	0.000	6	4	4	2
AB60304	Phone message: Write out correct message	310	0.895	0.462	0.000	1	5	2	3
AB81002	Consumer Reports books	311	0.975	0.570	0.000	4	3	5	2
AB20801	Bus schd: Take correct bus for given condition (2)	313	1.282	0.902	0.144	10	3	5	2
AB50402	Catalog order: Order product two	314	1.108	0.717	0.000	8	4	4	3
AB40401	Almanac: Find page containing chart for given info	314	0.771	0.397	0.000	5	4	3	2
AB21001	Bus schd: Take correct bus for given condition (4)	315	0.730	0.521	0.144	10	3	4	2
AB60502	Petroleum graph: Complete graph including axes	318	1.082	0.783	0.000	10	6	2	2
A120701	MasterCard/Visa statement	320	0.513	-0.015	0.000	6	2	4	2
AB20701	Bus schd: Take correct bus for given condition (1)	324	0.522	0.293	0.131	10	3	4	2
A131301	Tempra dosage chart	326	0.624	0.386	0.000	5	4	4	2
AB50501	Telephone bill: Mark information on bill	330	0.360	-0.512	0.000	7	4	4	2
AB91401	Consumer Reports index	330	0.852	0.801	0.000	7	3	5	3
AB30801	Almanac: Find page containing chart for given info	347	0.704	0.929	0.000	5	4	5	2
AB20901	Bus schd: After 2:35, how long til Flint&Acad bus	348	1.169	1.521	0.163	10	5	4	2
A130301	El Paso Gas & Electric bill	362	0.980	1.539	0.000	8	5	4	5
A120801	MasterCard/Visa statement	363	0.727	1.266	0.000	6	5	4	2
AB91301	Consumer Reports index	367	0.620	1.158	0.000	7	4	5	3
AB60501	Petroleum graph: Label axes of graph	378	1.103	1.938	0.000	11	7	2	5
AB30901	Almanac: Determine pattern in exports across years	380	0.299	0.000	0.000	7	5	5	3
A100701	Spotlight economy	381	0.746	1.636	0.000	10	5	5	2
A100501	Spotlight economy	386	0.982	1.993	0.000	10	5	5	5
A100401	Spotlight economy	406	0.489	1.545	0.000	10	5	5	2
AB51001	Income tax table	421	0.257	0.328	0.000	9	4	5	2
A100601	Spotlight economy	465	0.510	2.737	0.000	10	7	5	2

13.4.2.5 Validity evidence for the document scale

As with the prose tasks, an important piece of validation evidence concerns the document tasks newly developed for the 1990 literacy survey of job-seekers served by the Department of Labor. These new tasks were designed to reflect various aspects of the theoretical framework as it evolved from the 1985 young adult literacy assessment. Table 13-5 gives the results of regression analyses for the 1990 document literacy tasks as well as for the combined set (including the 1985 document tasks). Overall, the variance accounted for reached 92 percent for the new tasks and 87 percent for the combined set of the 1985 and 1990 document literacy tasks.

Table 13-5. Unstandardized regression coefficients and standard errors predicting RP80 task difficulties on the basis of four structure and process variables: 33 new document literacy tasks and 92 total document literacy tasks from the 1990 survey of job-seekers

	New Tasks			All Tasks		
	Coeff	StdErr	p	Coeff	StdErr	p
Structure Variable						
Structural complexity	5.17	1.91	.01	1.39	1.10	.21
Process Variables						
Type of match	24.12	3.70	.00	24.46	2.28	.00
Plausibility of distractors	23.84	4.11	.00	22.71	2.44	.00
Type of information	-1.35	4.93	.79	9.09	3.15	.00
Variance accounted for:						
R^2		92%			87%	
Adjusted R^2		91%			86%	
Degrees of freedom		28			87	

13.4.3 Quantitative Literacy

Since adults are often required to perform numerical operations in everyday life, the ability to perform quantitative tasks is an important area of adult literacy. To complete these types of tasks successfully, a respondent must perform arithmetic operations such as addition, subtraction, multiplication, or division either singly or in combination using numbers or quantities that are embedded in printed information.

At first glance, quantitative tasks might appear to represent fundamentally different skills from those involved in processing prose and documents. However, an analysis of tasks along this scale shows that the difficulty of these quantitative tasks is affected by the processing of the printed information in which they are contained and thus by the processing variables salient for prose and document tasks.

13.4.3.1 Quantitative variables

In general, it appears that many individuals can perform simple arithmetic operations when both the numbers and operations are made explicit. Yet, when these same operations are performed on numbers that must be located and extracted from different types of documents that contain similar but irrelevant information, or when these operations must be inferred from printed directions, quantitative tasks become increasingly difficult. To complete tasks on the quantitative scale, individuals are required to match information in a question or directive to information stated in one or more documents or pieces of text. In addition, tasks from the quantitative scale may require the reader to negotiate information that can serve as *plausible distractors* during the calculation of a correct response. Moreover, individuals are also required to process some type of printed information. While *type of information* varied for prose and documents, requested information for the quantitative tasks is always an amount. The stimulus materials for the quantitative tasks are mostly documents, and these vary widely in their structural complexity.

Thus, while the quantitative tasks include *structural complexity*, *type of match* and *plausibility of distractors* as defined for the prose and document tasks, they also involve two "formulate" variables that are unique to this scale. The first formulate variable, *operation specificity*, refers to the process of identifying (and sometimes entering) the numbers in an arithmetic expression, including the determination of the appropriate operation(s) that must be performed. Tasks tend to be more difficult when the numbers must be identified in a document and when these numbers are not in column format or adjacent to each other. Tasks also tend to become more difficult when the operation is not specified or when the wording in the question or directive does not contain an explicit semantic relation statement such as "how many" or "calculate the difference." This variable was coded from 1 (easiest) to 9 (most difficult) based on a set of additive rules reflecting the various facets stated here.

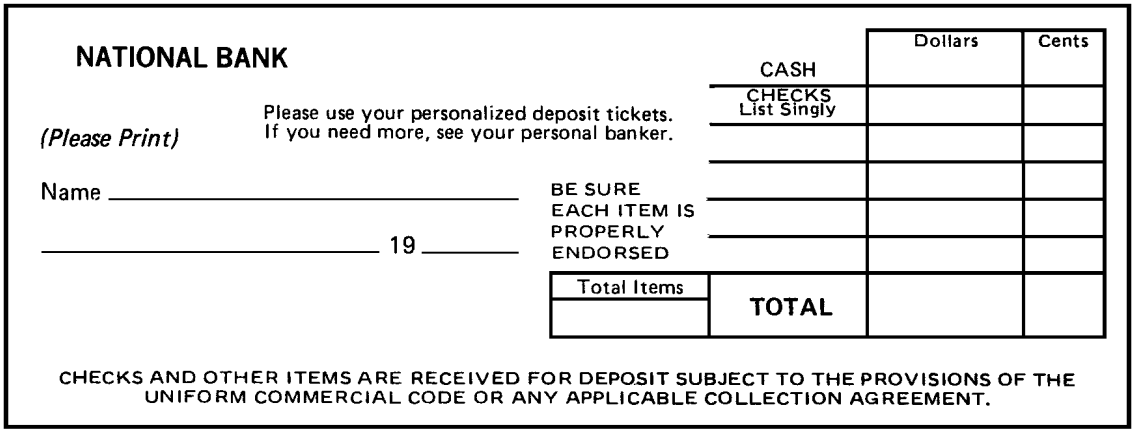
The second formulate variable, *type of calculation*, includes both the type of arithmetic operation (addition, subtraction, multiplication, or division) required to produce a correct response, and whether that operation must be performed alone or in combination. Tasks requiring two or more operations tend to be more difficult than those involving a single operation. This variable ranged from 1 (easiest) to 5 (most difficult).

13.4.3.2 Examples of quantitative literacy tasks

The least demanding quantitative task in the 1990 survey (220) required the reader to enter and total two numbers on a bank deposit slip (Exhibit 13-30). In this example, both the number and operation were judged to be easily identified, and the operation involved the simple addition of two decimal numbers that were presented in column format. Moreover, the numbers were stated in the directive so that the problem was, in some sense, set up for the reader. As a result, each of the process and formulate variables received a code of 1. The structural complexity of the document was coded 2.

In other tasks having similar characteristics that received somewhat higher values on the scale, the quantities, while easy to identify, were not explicitly given in the directive but had to be searched for and identified in the document. One such task having a scale value of 270 required the reader to locate the appropriate shipping charges in a table before entering the correct amount on an order form and calculating the total price for ordering office supplies. The structural complexity of this document was judged to be 6, while *type of match* was coded 3, and *plausibility of distractors* was coded 2. In addition, *type of calculation* received a code of 1 and *operation specificity* a code of 3.

Exhibit 13-30. Example of stimulus material for a relatively easy quantitative literacy task



NATIONAL BANK

(Please Print)

Please use your personalized deposit tickets. If you need more, see your personal banker.

Name _____

_____ 19 _____

BE SURE EACH ITEM IS PROPERLY ENDORSED

		Dollars	Cents
CASH			
CHECKS List Singly			
Total Items	TOTAL		

CHECKS AND OTHER ITEMS ARE RECEIVED FOR DEPOSIT SUBJECT TO THE PROVISIONS OF THE UNIFORM COMMERCIAL CODE OR ANY APPLICABLE COLLECTION AGREEMENT.

Tasks around 300 on the quantitative scale still require a single arithmetic operation. What appears to distinguish these tasks, however, is the fact that the reader must identify, in various places in the document, two or more numbers needed to solve the problem. The numbers are not presented in column format, nor is the operation needed to complete the task explicitly stated in the directive or provided by the format of the document, as in the previous examples. Instead, the operation must be determined from arithmetic relation terms, such as “how many” or “what is the difference” used in the question.

One such task receiving a scale value of 312 requires the reader to look at a table of money rates (Exhibit 13-31) to determine how much more interest would be earned in money market accounts provided by mutual funds than in those provided by S&Ls. This document received a structural complexity score of 4. It was also coded 3 for *type of match* and *plausibility of distractors* and 2 for each of the two formulate variables—*type of calculation* and *operation specificity*.

Tasks with the highest scale values (above 370) tended to require the reader to draw heavily on background information in order to identify both the quantities and the operations needed to complete the task successfully. For example, the most difficult quantitative task used in this assessment required readers to look at a newspaper advertisement for a home equity loan (Exhibit 13-32) and then, using the information provided, explain how they would calculate the total amount of interest charges to be paid. This document received a structural complexity score of 2. It was coded 5 for *type of match, plausibility of distractors*, and *type of calculation*, while *operation specificity* received a coded value of 7.

13.4.3.3 Coding the 1985 and 1990 Quantitative Literacy Tasks

The preceding section provided examples to illustrate how specificity of operation, type of calculation, and plausibility of distractors were coded for selected quantitative literacy tasks from the 1990 literacy survey of job-seekers served by the Department of Labor. In this section, the coding rules are formulated

Exhibit 13-31. Example of stimulus material for relatively difficult quantitative literacy task

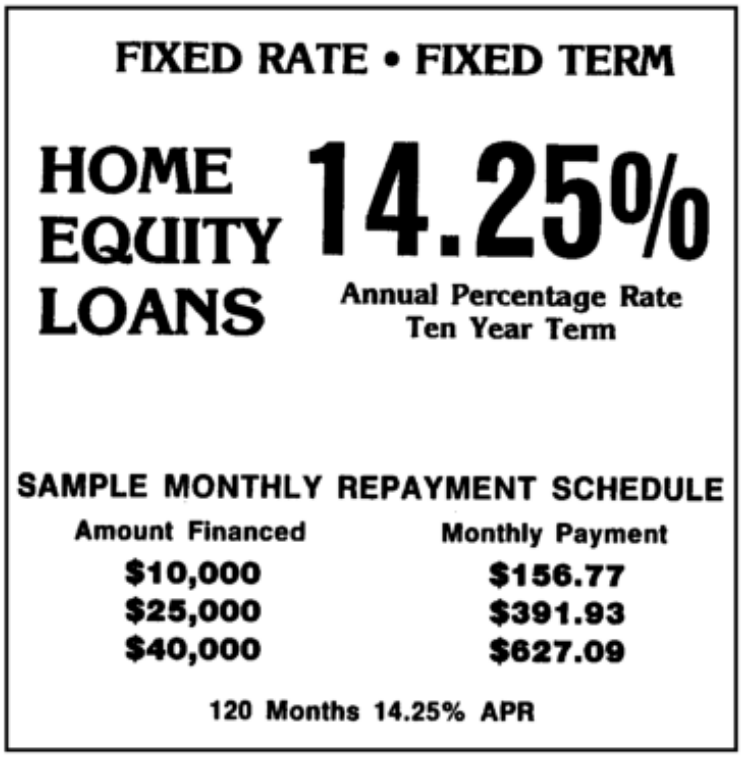
MONEY RATES

	Thurs.	6 mo. ago	Yr. ago
Prime lending	10.00%	8.50%	8.75%
Fed discount	6.50%	6.00%	6.00%
Broker call loan	9.13%	7.63%	8.13%
Mortgage rates			
30-yr. fixed-rate (FHLMC)	10.65%	9.85%	10.63%
30-yr. adjustable (FHLMC)	8.16%	7.53%	7.84%
15-vr. fixed rate[1]	10.39%	9.75%	10.28%
ARM index (1-year Treas.)	8.24%[2]	6.63%	7.41%
Money market accounts, latest 7-day average			
Money mutual funds[3]	7.37%	6.05%	6.03%
Banks and S&Ls[4]	5.81%	5.59%	5.47%
Treasury security rates			
3-month T-bill discount[4]	7.26%	5.74%	6.45%
6-month T-bill discount[4]	7.40%	5.93%	6.72%
7-year note	8.85%,-.01	8.12%	9.22%
30-year bond	9.03%, -.03	8.55%	9.57%

1—*Bank Rate Monitor*
2—week ending Sept. 2
3—*Donoghue's Money Fund Report*
4—Sept. 6 auction

THE DOLLAR

Exhibit 13-32. Example of stimulus material for a difficult quantitative literacy task



in more general terms. A number of criteria must be taken into account when measuring the four variables associated with task difficulty on the quantitative literacy scale.

13.4.3.3.1 Specificity of Operation

Description. To obtain the requested quantitative information, readers must identify the relevant quantities, understand their relationships to one another, set up an arithmetic equation based on these relationships, and carry out simple arithmetic manipulations. Instructions for how to formulate an arithmetic expression can be more or less specific about identifying the relevant quantities and the relationships among them. Further, the operations necessary to obtain the requested information can be more or less specific.

Scoring rules. The evaluation of operation specificity takes into account aspects of both the necessary operation and the amounts involved. The *relationship* among the relevant amounts can be specified in terms that directly translate into an operation, that translate more indirectly, or that require a larger quantitative vocabulary to understand the relationship. Quantitative tasks are easier if the relationships among the amounts are described with arithmetic symbols or common arithmetic terms. Tasks are more difficult if the terminology describing the relationship uses a more specialized vocabulary or requires understanding ratios.

Exhibit 13-33. Basic scoring rule for specificity of required operation: Relationships of amounts

Rule	Score
When the operation is specifically identified by words or symbols, using terms such as ‘add’, ‘total’ or ‘+’; ‘subtract’ or ‘-’; ‘multiply’ or ‘x’; and ‘divide’ or ‘÷’	0
When the operation is specified with a relationship among the numbers, using terms such as ‘how much more’; ‘how much less’ or ‘calculate the difference’; and ‘how many times’	1
When the operation is identified using more specialized vocabulary, using terms such as ‘how much is saved’; ‘how much is the deduction’; or ‘what is the net profit’	2
When the operation is identified as a unit ratio, such as ‘miles per gallon’, ‘cost per square foot’, or ‘price per square yard’	3

Operation specificity is made more difficult as the *amounts* involved are harder to identify or harder to use in arithmetic operations. Quantitative tasks are easier if the amounts appear in a row or column format, if they are adjacent to one another, if they are labeled, if they do not require a search, if they are one-step problems, or if they do not involve conversions of units of measurement. Tasks are harder if they are not in a row or column format, if the amounts are not adjacent, if the labels associated with the amounts have to be inferred, if the amounts require a search, if they involve more than one step, or if they require converting units of measurement.

Exhibit 13-34. Additional scoring rule for specificity of required operation: Identifying amounts

Rule	Add
If the amounts are • in a row and column format • adjacent to one another • presented in the current task and no search is needed, if the problem requires a single step, if the labels are present and no inference is needed, or if the amounts need not be transformed into common units	0
If the amounts are *not* in a row and column format	1
If the amounts are *not* adjacent to one another	1
If the amounts are *not* presented, but must be carried over from a prior task, or the problem requires more than one step	1
If the amounts are *not* presented, but must be identified by a search	1
If the labels for the amounts must be inferred	1
If the amounts are in different units (such as time in hours and minutes or fractions with different denominators) that must be transformed into a common unit	1

The scoring rules for operation specificity are additive and scored by adding a point to the basic score for any of several possible factors that can make the relevant amounts more difficult to identify and manipulate. A quantitative task, for example, might have a basic score of 1 because it specifies with common terminology a relationship among the numbers, but has additional points added because the amounts are not in a row and column format (add 1) nor are they adjacent to one another (add 1), and the amounts must be transformed into a common unit of measurement (add 1). A quantitative assessment task with these features would have a total operation-specificity score of 4.

13.4.3.3.2 Type of Calculation

Description. Type of calculation measures the complexity of the various operations that readers use to relate one set of numbers to another in order to produce a sum, difference, product, or quotient. Quantitative tasks are easiest when the calculation is a single sum and most difficult when more than one of these calculations is required.

Scoring rules. For tasks that involve a single operation, those that involve addition are the easiest; those that involve subtraction are next easiest; those that involve multiplication are more difficult; and those involve division are the most difficult. When the reader must manipulate numbers that are the outcome of operations in preceding tasks, then the job becomes even more difficult. Any task that requires two or more operations (such as a division followed by a multiplication) is more difficult than any that requires only a single operation.

Exhibit 13-35. Scoring rules for type of calculation

Rule	Score
When two quantities are to be added (easiest)	1
When two quantities are to be subtracted	2
When two quantities are to be multiplied	3
When one quantity is to be divided by another	4
When readers are requested to operate on two quantities and use the result with another quantity to perform a second operation (hardest)	5

13.4.3.3.3 Type of Match, Plausibility of Distractors, and Structural Complexity

The task features type of match and plausibility of distractors apply to quantitative literacy tasks in the same way as they did to prose and document literacy tasks. Structural complexity is a feature of the prose text or documents in which the quantitative information is embedded. It was also defined in the same way as it was for the prose and document literacy tasks.

13.4.3.4 Codes for all 1985 and 1990 quantitative literacy tasks

The preceding sections described in detail the several criteria that must be taken into account when measuring the variables associated with task difficulty on the quantitative literacy scale. These rules were applied to all quantitative literacy tasks in the 1990 survey and in the 1985 young adult literacy assessment. The resulting codes, along with RP80 task difficulties and IRT item parameters are shown in Table 13-6.

Table 13-6. List of quantitative literacy tasks, along with RP80 task difficulty, IRT item parameters, and values of variables associated with task difficulty (structural complexity, type of match, plausibility of distractors, type of calculation, and specificity of operation): 1990 survey of the literacy of job-seekers

Identifier	Quantitative Literacy Items	RP80	IRT parameters			Complexity	Match	Distractor	Calculation	Op specfy
			a	b	c					
AB70904	Enter total amount of both checks being deposited	221	0.869	-1.970	0.000	2	1	1	1	1
AB50404	Catalog order: Shipping, handling, and total	271	0.968	-0.952	0.000	6	3	2	1	3
AB91201	Tempra coupon	271	0.947	-0.977	0.000	1	2	1	5	4
AB40701	Check ledger: Complete ledger (1)	277	1.597	-0.501	0.000	3	2	2	1	4
A121001	Insurance protection workform	275	0.936	-0.898	0.000	2	3	2	3	2
AB90102	Pest control warning	279	0.883	-0.881	0.000	2	3	3	1	4
AB40702	Check ledger: Complete ledger (2)	281	1.936	-0.345	0.000	3	2	2	2	4
AB40703	Check ledger: Complete ledger (3)	282	1.874	-0.332	0.000	3	1	2	2	4
A131601	Money rates: Thursday vs. one year ago	281	1.073	-0.679	0.000	4	3	2	2	4
AB40704	Check ledger: Complete ledger (4)	283	1.970	-0.295	0.000	3	2	2	2	4
AB80201	Burning out of control	286	0.848	-0.790	0.000	2	3	2	2	4
A110101	Dessert recipes	289	0.813	-0.775	0.000	5	3	2	2	4
AB90201	LPGA money leaders	294	0.896	-0.588	0.000	5	2	2	2	4
A120101	Businessland printer stand	300	1.022	-0.369	0.000	2	3	3	2	4
AB81003	Consumer Reports books	301	0.769	-0.609	0.000	7	2	3	1	4
AB80601	Valet airport parking discount	307	0.567	-0.886	0.000	2	3	3	2	4
AB40301	Unit price: Mark economical brand	311	0.816	0.217	0.448	2	2	3	4	6
A131701	Money rates: compare S&L w/ Mutual funds	312	1.001	-0.169	0.000	4	3	3	2	2
AB80701	Valet airport parking discount	315	0.705	-0.450	0.000	2	2	3	3	4

Table 13-6. List of quantitative literacy tasks, along with RP80 task difficulty, IRT item parameters, and values of variables associated with task difficulty (structural complexity, type of match, plausibility of distractors, type of calculation, and specificity of operation): 1990 survey of the literacy of job-seekers — Continued

Identifier	Quantitative Literacy Items	RP80	IRT parameters			Complexity	Match	Distractor	Calculation	Op specfy
			a	b	c					
A100101	Pizza coupons	316	0.690	-0.472	0.000	2	3	3	1	4
AB90301	LPGA money leaders	320	1.044	0.017	0.000	5	1	2	4	3
A110401	Dessert recipes	323	1.180	0.157	0.000	5	3	2	3	6
A131401	Tempra dosage chart	322	1.038	0.046	0.000	5	3	3	2	4
AB40501	Airline schedule: plan travel arrangements (1)	326	0.910	0.006	0.000	3	3	3	5	3
AB70501	Lunch: Determine correct change using info in menu	331	0.894	0.091	0.000	2	2	2	5	4
A120201	Businessland printer stand	340	0.871	0.232	0.000	2	3	4	3	5
A110901	Washington/Boston train schedule	340	1.038	0.371	0.000	7	4	4	2	5
AB60901	Nurses Convention: Write number of seats needed	346	0.504	-0.355	0.000	3	4	4	1	5
AB70601	Lunch: Determine 10% tip using given info	349	0.873	0.384	0.000	2	1	2	5	7
A111001	Washington/Boston train schedule	355	0.815	0.434	0.000	7	4	4	2	5
A130501	El Paso Gas & Electric bill	352	0.772	0.323	0.000	8	3	4	2	2
A100801	Spotlight economy	356	0.874	0.520	0.000	8	5	4	2	2
AB40201	Unit price: Estimate cost/oz of peanut butter	356	0.818	0.455	0.000	2	1	2	4	5
A121101	Insurance protection workform	356	0.860	0.513	0.000	2	1	2	5	4
A100901	Camp advertisement	366	0.683	0.447	0.000	2	2	4	5	4
A101001	Camp advertisement	366	0.974	0.795	0.000	2	3	4	5	4
AB80501	How companies share market	371	1.163	1.027	0.000	6	3	2	3	6
A131501	Tempra dosage chart	381	0.916	1.031	0.000	5	3	5	3	5
AB50403	Catalog order: Order product three	382	0.609	0.601	0.000	6	4	5	5	5
AB91001	U.S. Savings Bonds	385	0.908	1.083	0.000	6	4	5	2	4
A110601	Registration & tuition Info	407	0.624	1.078	0.000	8	2	5	5	5
AB50301	Interest charges: Orally explain computation	433	0.602	1.523	0.000	2	5	5	5	7

13.4.3.5 Validity evidence for the quantitative scale

As with the prose tasks, one piece of validation evidence concerns the quantitative tasks newly developed for the 1990 survey. There were only 15 quantitative tasks in the 1985 young adult literacy assessment; to fill in the scale, 28 new tasks were developed for the 1990 assessment using the theoretical framework described here. As shown in Table 13-7, the combined set of structural complexity, process, and formulate variables accounts for 84 percent of the variance in scale values for the 28 new 1990 tasks and 83 percent of the variance for the combined set of 43 tasks from the combined task set from the 1985 and 1990 assessments.

Table 13-7. Unstandardized regression coefficients and standard errors predicting RP80 task difficulties on the basis of five structure, process, and formulate variables: 28 new quantitative literacy tasks and 43 total quantitative literacy tasks from the 1990 survey of job-seekers

	New Tasks			All Tasks		
	Coeff	StdErr	p	Coeff	StdErr	p
Structure Variable						
Structural complexity	4.81	1.60	.01	4.11	1.59	.01
Process Variables						
Type of match	1.25	4.58	.79	.06	3.82	.99
Plausibility of distractors	20.44	3.52	.00	21.21	3.69	.00
Formulate Variables						
Type of calculation	11.56	3.07	.00	10.76	2.25	.00
Operation specificity	8.23	2.76	.01	9.57	2.44	.00
Variance accounted for:						
R^2		84%			83%	
Adjusted R^2		81%			81%	
Degrees of freedom		22			37	

13.4.4 Establishing Proficiency Levels

As the public report for the 1990 assessment of job-seekers was being drafted, the need to clarify the meaning of the scales as well as the graphic presentations of the results became increasingly clear. The challenge was to find a way to avoid the information overload of the figures and tables that were produced for the 1985 young adult literacy assessment (see Exhibit 13-2) and to pass along our growing understanding of appropriate score interpretation.

Empirical data of the kind presented in this chapter for each of the three literacy scales provides evidence for the credibility of the notion that while literacy is not a single skill suited to all types of tasks, neither is it an infinite number of skills each associated with a given prose, document, or quantitative task. Rather, as the data presented here indicate, there appears to be an ordered set of information-processing skills and strategies that may be called into play to accomplish the range of tasks represented by the three literacy domains.

As the project team sought effective ways to present the data from the 1990 survey, they re-examined the implications of the increases in process complexity as task difficulty rises. As tasks became more difficult, their associated code values on the task variables also increased (See Tables 13-2, 13-4, and 13-6). This relationship between task difficulty and code values appeared to be quite systematic. That is, toward the bottom of each literacy scale, the code value of 1 on each task variable was dominant; values of 2 and 3 became more frequent as tasks moved up the prose, document, and quantitative scales;

and toward the higher end, code values of 4, 5, and higher become predominant. Although the patterns differed somewhat from scale to scale reflecting differences in the coded values assigned to the variables, major shifts in the processes and skills required for successful task performance occurred at remarkably similar points on the three scales.

Visual inspection of the distributions of task codes along each of the literacy scales revealed several major points occurring at roughly 50 point intervals beginning with 225 on each scale. As with all systems, this one contains some noise and does not account for all of the score variance associated with performance on the three literacy scales. Moreover, the shifts in skill or process requirements do not necessarily occur at exactly 50 point intervals on the scales. However, assigning the exact range of scores to capture each level (for example, using score 277-319 to represent Level 3 on the document scale and 331-370 to represent Level 4 on the quantitative scale) implies a precision of measurement that is inappropriate for the methodology adopted. In order to ensure consistency across the scales, 50 point intervals were imposed. The advantage of having common intervals outweighs the marginal gain in agreement percentages and offers better readability to the users. Consequently, on the basis of the distributions of processing requirements, the results of the 1990 survey were aggregated into five proficiency levels depending on the range of the scores: Level 1 (less than or equal to 225), Level 2 (from 226 to 275), Level 3 (from 276 to 325), Level 4 (from 326 to 375), and Level 5 (greater than or equal to 376).

Once the levels were tentatively set, criteria to account for task placement within levels were determined, based solely on inspection of the code values assigned to each task. These criteria and the percentages of tasks meeting these criteria are shown in Table 13-8. Overall, an average of 78 percent of the prose tasks met the identified criteria for each level. The agreement percentages on the document and quantitative scales were 89 and 79, respectively. The advantage of having common intervals across scales outweighs the marginal gain in agreement percentages, thus were implemented.

Table 13-8. Numerical criteria assigned to task variables to distinguish proficiency levels and percentages of tasks meeting those criteria for prose, document, and quantitative literacy scales

	Prose literacy		Document literacy		Quantitative literacy	
	Criteria*	Percent agreement	Criteria[H]	Percent agreement	Criteria[I]	Percent agreement
Level 1	1, 1, ≤2	100	1, ≤2, ≤2	87	1, 1, ≤2	100
Level 2	2, 2, 2 or 3, ≤3, ≤3	73	2, 2, 2 or 3, ≤3, ≤2	97	3, ≤2, ≤2	50
Level 3	4, ≤3, ≤3	80	≥4, ≤3, ≤3	87	≤5, ≤3, ≤3	82
Level 4	4, ≥4, ≤4	73	4, ≥4, ≤3	75	≥4, ≤4, ≤4	77
Level 5	5, ≥5, ≤5	75	5, ≥5, ≤5	86	5, ≥5, ≤5	80
Overall agreement		78		89		79

*Criteria for prose literacy tasks pertain to type of match, plausibility of distractors, and type of information. [H]Criteria for document literacy tasks pertain to type of match, plausibility of distractors, and type of information. [I]Criteria for quantitative literacy tasks pertain to operation specificity, type of calculation, and plausibility of distractors.

The next step in evaluating the utility of using the five identified levels as reporting categories was to run regression analyses using levels rather than individual task scale values as the dependent variable. These results are presented in Table 13-9. As shown here, the identical process variables are significant in predicting proficiency levels as was the case in predicting specific task scale values. Moreover, the models used to predict proficiency levels account for roughly the same amount of variance as those used to predict task values—ranging from 78 percent on the quantitative scale to 80 percent on the prose scale to 88 percent on the document scale. These data are somewhat surprising given the typical effects of restriction of range on correlational data.

Table 13-9. Unstandardized regression coefficients and standard errors predicting five levels of RP80 task difficulty on the basis of six structure, process, and formulate variables: All prose, document, and quantitative literacy tasks from the 1990 survey of job-seekers

	Prose literacy			Document literacy			Quantitative literacy		
	Coeff	StdErr	p	Coeff	StdErr	p	Coeff	StdErr	p
Structure Variable									
Readability/structural complexity	.03	.04	.53	.00	.20	.93	.07	.37	.07
Process Variables									
Type of match	.19	.08	.03	.46	.04	.00	-.05	.09	.58
Plausibility of distractors	.49	.09	.00	.43	.04	.00	.48	.09	.00
Type of information	.25	.08	.01	.21	.06	.00	-	-	-
Formulate Variables									
Type of calculation	-	-	-	-	-	-	.14	.05	.01
Operation specificity	-	-	-	-	-	-	.20	.06	.00
Variance accounted for:									
R^2	80%			88%			78%		
Adjusted R^2	78%			87%			75%		

In addition to accounting for significant amounts of variance, the variables described in this chapter illustrate the internal consistency of processing characteristics associated within each of the five levels. This, in turn, provides evidence of the substantive aspect of construct validity, or the appropriateness of the theoretical model predicted to underlie consistencies in performance. Given this evidence concerning the validity of these five levels, it was decided to use them to report the distributions of the literacy of job-seeking adults participating in the two Department of Labor programs (the Job Training Partnership Act and the U.S. Employment Service) and to communicate the meaning of what was being measured along each of the scales.

Brief statements were developed to summarize some of the knowledge and skills associated with successful performance within each of the identified levels. These descriptions were derived from the pattern of codes among the processing variables associated with tasks falling within a level and are shown in Table 13-10. Displaying tasks along each scale as was done for the 1985 young adult literacy assessment (see Exhibit 13-36) tends to encourage interpretation of the scale at the task level. Rather than simply displaying tasks along each scale, the use of the five levels allowed the development of descriptive information that could be combined with the percentages of adults in various subpopulations who demonstrated performance within each of the designated levels.

Table 13-10. Descriptions of prose, document, and quantitative literacy levels based on type of match, plausibility of distractors, type of information, operation specificity, and type of calculation: 1990 survey of job-seekers

Levels	Prose literacy	Document literacy	Quantitative literacy
Level 1 0-225	Prose tasks at this level are the least demanding in terms of what the reader must do to produce a correct response. Typically, tasks at this level require the reader to locate one piece of information in which there is a literal match between the question and the stimulus material. If a distractor or plausible answer is present, it tends to be located away from where the correct information is found.	Tasks at this level are the least demanding. In general, they require the reader to either locate a piece of information based on a literal match or to enter information from personal knowledge.	Although no quantitative tasks used in this assessment fall within this level, experience suggests that such tasks would require a single, relatively simple operation for which the numbers are given and the operation specified.
Level 2 226-275	Some of the prose tasks of this level still require the reader to locate on a single literal feature of information; however, these tasks tend to occur in materials where there are several distractors or where the match is based on low-level inferences. Tasks at this level also begin to require the readers to integrate information by pulling together two or more pieces of information or by comparing and contrasting information.	Tasks at this level begin to become more varied. Some still require the reader to match a single match a single piece of information; however, tasks occur where there are several distractors or where the match is based on low-level inferences. Tasks at this level also begin to require the reader to cycle through information or to integrate information.	Tasks at this level typically require the use of a single operation based on numbers that are either stated in the question or easily located in the material. In addition, the operation needed is either stated in the question or easily determined based on the format of the problem—for example, entries on a bank deposit slip or order form.
Level 3 276-325	Tasks at this level tend to require the reader to search fairly dense text for literal or synonymous matches on the basis of more than one feature of information or to integrate information from relatively long text that does not contain organizational aids such as headings.	Tasks at this level tend to require the reader to either integrate three pieces of information or to cycle through materials in rather complex tables or graphs in which distractor information is present.	What appears to distinguish tasks at this level is that two or more numbers needed to solve the problem must be found in the stimulus material. Also the operation(s) needed can be determined from arithmetic relation terms.
Level 4 326-375	Tasks at this level continue to demand more from the reader. Not only are multiple-feature matching and integration of information from complex displays materials maintained, the degree of inferencing required by the reader is also increased. Conditional information is frequently present in tasks at this level that must be taken into account.	Tasks at this level tend to demand more from the reader. Not only are multiple-feature matching, cycling, and integration of information maintained, the degree of inferencing is increased. Cycling tasks often require the reader to make five or more responses with no designation of the correct number of responses. Conditional information is also present and must be taken into account.	Quantitative tasks at level 4 tend to require two or more sequential operations or the application of a single operation where either the quantities must be located in complex displays and/or the operation must be inferred from semantic information given or prior knowledge.
Level 5 376-500	These tasks require the reader to search for information in dense text or complex documents containing multiple plausible distractors, to make high text-based inferences or use specialized background knowledge, as well as to compare and contrast sometimes complex information to determine differences.	Tasks at this level require the most from the reader. The reader must search through complex displays contain[ing] multiple distractors, make high text-based inferences or use specialized knowledge.	Quantitative tasks at this level are the most demanding. They tend to require the reader to perform multiple operations and to disembed features of a problem from stimulus material or to rely on background knowledge to determine the quantities or operations needed.

In addition, using information derived from the IRT analyses, it became possible to estimate the likelihood that individuals with various proficiency levels would perform the average task within a specified level correctly. These distributions of likelihood estimates provide a richer and more accurate reflection of the range of tasks that an individual can be expected to perform successfully. Collectively, the descriptors, distributions, and probabilities of correctly performing various tasks within different levels could be displayed in a single table. An example, taken from the 1990 public report (Kirsch, Jungeblut and Campbell, 1992), is reproduced as Exhibit 13-36.

Exhibit 13-36. Descriptions of five document literacy levels, average RP80 scale values, and probabilities of performing tasks in the level successfully: Adult job seekers, 1990

Levels	Description of Document Tasks at Each of Five Levels	Average RP80 at Each Level	Average Probability at Selected Proficiency Levels						TOTAL
			200	250	300	350	400		
Level 1 0I-225	Tasks at this level are the least demanding. In general, they require the reader to either locate a piece of information based on a literal match or to enter information from personal knowledge.	194	80	94	98	100	100	JTPA ES/UI Young Adults	14.1 (2.0) 13.1 (1.6) 8.0 (0.6)
Level 2 226-275	Tasks at this level begin to become more varied. Some still require the reader to math a single piece of information; however, tasks occur where there are several distractors or where the match is based on low-level inferences. Tasks at this level also begin to require the reader to cycle through information or to integrate information.	248	47	81	95	99	100	JTPA ES/UI Young Adults	37.3(1.3) 30.1 (1.2) 24.2 (1.1)
Level 3 276-325	Tasks at this level tend to require the reader to either integrate three pieces of information or to cycle through materials in rather complex tables or graphs in which distractor information is present.	300	30	54	79	93	97	JTPA ES/UI Young Adults	35.4 (1.5) 35.9 (1.0) 39.7(1.2)
Level 4 326-375	Task at this level continue to demand more from the reader. Not only are multiple-feature matching, cycling, and integration of information maintained, the degree of inferencing is increased. Cycling tasks often require the reader to make five or more responses with no designation of the correct number of responses. Conditional information is also present and must be taken into account.	351	11	26	53	79	93	JTPA ES/UI Young Adults	12.2 (1.8) 18.5 (1.7) 24.0 (1.1)
Level 5 376-500	Tasks at this level require the most from the reader. The reader must search through complex displays containing multiple distractors, make high text-based inferences, or use specialized knowledge.	405	15	23	37	60	79	JTPA ES/UI Young Adults	1.1 (0.4) 2.4 (0.5) 4.1 (0.6)

*The numbers in parentheses are estimated standard errors.

It will be seen, for example, that about 37 percent of the JTPA participants demonstrated performance in the Level 2 range of the document scale. The average difficulty of tasks in this level is 248, and an individual scoring at 250 has a probability of 81 percent of performing such a task successfully. Individuals scoring at 300 and above are likely to make few errors on tasks at around 248 on the document scale. Similarly, an individual with a proficiency score of 250 has a better than 90 percent chance of responding correctly to tasks in Level 1. This same individual has a probability of about 50 percent of successfully performing Level 3 tasks and about a 25 percent probability of performing Level 4 and Level 5 tasks correctly.

13.5 THE 1992 NATIONAL ADULT LITERACY SURVEY

The 1992 National Adult Literacy Survey assessed the literacy skills of a nationally representative sample of individuals age 16 and older, as well as representative samples of individuals ages 16-64, in 12 states (Kirsch, Jungeblut, Jenkins, and Kolstad, 1993). The National Center for Education Statistics requested

that the assessment results be linked to both the 1985 young adult literacy assessment and the 1990 survey of adult job-seekers served by the U.S. Department of Labor. To help meet the survey's objectives, a Literacy Definition and a Technical Review Committee were established. After some deliberations, members of the Literacy Definition Committee recommended the adoption of the same definition and measurement framework used in the two earlier surveys. In addition, members of this committee also requested that the results from the 1992 National Adult Literacy Survey be reported in terms of the same five levels developed for the 1990 survey of job-seekers. This decision provided a unique opportunity to investigate further the validity of the theoretical framework that had evolved for developing new assessment tasks and for enhancing score meaning.

13.5.1 Prose Literacy

The development of the 1992 National Adult Literacy Survey entailed both the reuse of existing prose literacy tasks from the 1985 young adult literacy assessment and the production of new prose literacy tasks. The new tasks continued the emphasis on expository prose drawn from authentic sources that adults might ordinarily encounter in daily life. The resulting assessment pool for the 1992 National Adult Literacy Survey included 41 prose literacy tasks of which 27 were newly developed for the 1992 survey.

A preceding section described in detail the several criteria that must be taken into account when measuring the four variables associated with task difficulty on the prose scale. Just as with the 1990 survey and the 1985 young adult literacy assessment, these rules were applied to all prose literacy tasks in the 1992 National Adult Literacy Survey, and the resulting codes, along with RP80 task difficulties and IRT item parameters are shown in Table 13-11.

Table 13-11. List of new prose literacy tasks, along with RP80 task difficulty, IRT item parameters, and values of variables associated with task difficulty: The 1992 National Adult Literacy Survey

Identifier	Task Description	Scaled RP80	IRT parameters			Read-ability	Type of match	Distractor Plausibility	Information type
			a	b	c				
NC00301	"My Dream:" Find country in short story	150	0.893	-3.228	0.000	4	1	1	1
N120901	Susan Butcher: Find number of wins of sled race	210	0.889	-2.061	0.000	9	1	1	2
NC00401	"My Dream:" Underline sentence explaining action	224	0.765	-1.936	0.000	4	1	2	4
N080101	SSI: Mark correct information in article	226	1.329	-1.447	0.000	6	1	1	3
N100101	"Growing Up:" Find first buyer's name	239	1.467	-1.147	0.000	8	3	2	1
N090601	Face off: What group will mandate safe cars?	253	1.878	-0.748	0.000	10	3	2	1
N090701	Face off: Find correct information in article	256	1.805	-0.699	0.000	10	3	2	2
N110101	Blood pressure: Why difficult to know if high	262	0.988	-0.971	0.000	7	3	2	4
N130801	Cost to raise child: Find information from article	274	0.735	-1.013	0.000	6	2	4	2
N110501	Jury: Underline sentence explaining action	276	0.939	-0.731	0.000	7	4	3	3
N080201	SSI: What must an SSI user accept if offered?	277	1.516	-0.389	0.000	6	4	2	3
N100201	"Growing Up:" Determine correct day of delivery	284	1.297	-0.346	0.000	8	4	3	2
N100301	"Growing Up:" What reason given to stop selling?	287	1.187	-0.344	0.000	8	5	1	4
N010201	Marketing: Underline sentence explaining action	288	1.059	-0.403	0.000	15	3	4	3
N110401	Jury: Length of time served by a juror	314	0.770	-0.192	0.000	7	4	2	4
N120301	Ida Chen: What experience turned Ida toward law?	316	1.075	0.142	0.000	7	4	2	3
N120401	Two things Chen did to resolve discrimination conflicts	317	1.162	0.229	0.000	7	4	3	2
N130201	"Fueled:" Determine phrase meaning	324	1.089	0.316	0.000	9	5	1	3
N130401	"Fueled:" Give suggestion about good value change	346	1.576	0.979	0.000	9	5	1	4
N010101	Marketing: List two facts	349	0.869	0.608	0.000	15	5	5	4
N090801	Contrast views on fuel-efficiency vs. Size of car	360	1.239	1.091	0.000	10	6	2	5
N080301	SSI: What is most you can make to receive SSI?	362	0.619	0.486	0.000	6	4	5	2
N130301	"Fueled:" Give diff and similarity between events	375	0.978	1.214	0.000	9	6	2	4
N100401	"Growing Up:" Compare approaches to selling mags	383	0.842	1.236	0.000	8	6	2	5
N110601	Two challenges attorneys use to jurors	410	1.045	1.954	0.000	6	6	2	5
N120501	Ida Chen: Interpret phrase from article	424	0.927	2.107	0.000	7	6	3	5
N010301	Marketing: Give purpose of event	433	0.787	2.138	0.000	15	5	5	3

Another piece of validation evidence can be obtained from analysis of the tasks used in the 1992 National Adult Literacy Survey. The four task variables were used in a regression analysis designed to predict the RP80 scale values of each task. Table 13-12 shows the results of these analyses for both the new tasks as well as the entire 1992 pool of prose literacy tasks. The numbers in the tables represent the raw beta coefficients for each of the variables included in the regression analyses along with standard errors and probabilities. Overall, the three process variables were significant for both the new 1992 tasks and for the complete set of prose tasks. The same variables as were found with the 1990 prose literacy tasks contribute to the predictive models. In addition, the amount of variance accounted for in the new tasks is similar—81 percent on the 1990 survey and 89 percent on the 1992 survey.

Table 13-12. Unstandardized regression coefficients and standard errors predicting prose literacy RP80 task difficulties on the basis of four structure and process variables: 27 new prose literacy tasks and 41 total prose literacy tasks from the 1992 survey of adults.

Prose literacy	New 1992 Tasks			All 1992 Tasks		
	Coeff	StdErr	p	Coeff	StdErr	p
Structure Variable						
Readability	-.07	1.80	.68	.18	1.67	.69
Process Variables						
Type of match	29.65	3.62	.00	28.86	3.45	.00
Plausibility of distractors	18.41	4.06	.00	16.09	3.63	.00
Type of information	10.91	4.35	.02	8.84	4.17	.04
Variance accounted for:						
R^2		89%			87%	
Adjusted R^2		87%			86%	
Degrees of freedom		22			36	

Since the equations predicting prose literacy task difficulty in the 1992 data essentially reproduced the findings of the 1990 data, and since there was a contractual requirement for comparability with the prior surveys, there was no need to revisit the cutpoints for the literacy levels. Still, the 1992 data provided an occasion to improve the language describing the literacy levels in minor ways. Table 13-13 provides a comparison of verbal descriptions of the prose literacy levels used in reporting the 1990 and 1992 survey results. The minor adjustments that can be seen in the 1992 descriptions were designed to make them more consistent with the variables predicting task difficulty.

Table 13-13. Descriptions of prose literacy levels based on type of match, plausibility of distractors, and type of information: 1990 survey of job-seekers and the 1992 National Adult Literacy Survey.

Levels	1990 Description	1992 Description
Level 1 0-225	Prose tasks at this level are the least demanding in terms of what the reader must do to produce a correct response. Typically, tasks at this level require the reader to locate one piece of information in which there is a literal match between the question and the stimulus material. If a distractor or plausible answer is present, it tends to be located away from where the correct information is found.	Most of the tasks in this level require the reader to read relatively short text to locate a single piece of information which is identical to or synonymous with the information given in the question or directive. If plausible but incorrect information is present in the text, it tends not to be located near the correct information.
Level 2 226-275	Some of the prose tasks of this level still require the reader to locate on a single literal feature of information; however, these tasks tend to occur in materials where there are several distractors or where the match is based on low-level inferences. Tasks at this level also begin to require the readers to integrate information by pulling together two or more pieces of information or by comparing and contrasting information.	Some tasks in this level require readers to locate a single piece of information in the text; however, several distractors or plausible but incorrect pieces of information may be present, or low-level inferences may be required. Other tasks require the reader to integrate two or more pieces of information or to compare and contrast easily identifiable information based on a criterion provided in the question or directive.
Level 3 276-325	Tasks at this level tend to require the reader to search fairly dense text for literal or synonymous matches on the basis of more than one feature of information or to integrate information from relatively long text that does not contain organizational aids such as headings.	Tasks in this level tend to require readers to make literal or synonymous matches between the text and the information given in the task, or to make matches that require 1ow-level inferences. Other tasks ask readers to integrate information from dense or lengthy text that contains no organizational aids such as headings. Readers may be asked to generate a response based on information that can be easily identified in the text. Distracting information is present, but is not located near the correct information.
Level 4 326-375	Tasks at this level continue to demand more from the reader. Not only are multiple-feature matching and integration of information from complex displays materials maintained, the degree of inferencing required by the reader is also increased. Conditional information is frequently present in tasks at this level that must be taken into account.	These tasks require readers to perform multiple-feature matches and to integrate or synthesize information from complex or lengthy passages. More complex inferences are needed to perform successfully. Conditional information is frequently present in tasks at this level and must be taken into consideration by the reader.
Level 5 376-500	These tasks require the reader to search for information in dense text or complex documents containing multiple plausible distractors, to make high text-based inferences or use specialized back-ground knowledge, as well as to compare and contrast sometimes complex information to determine differences.	Some tasks in this level require the reader to search for information in dense text which contains a number of plausible distractors. Others ask readers to make high-level inferences or use specialized background knowledge. Some tasks ask readers to contrast complex information.

13.5.2 Document Literacy

Just as with the prose literacy scale, the development of the 1992 National Adult Literacy Survey entailed both the reuse of existing document literacy tasks from the 1985 young adult literacy assessment and the production of new document literacy tasks. The new tasks continued the emphasis on documents drawn from authentic sources that adults might ordinarily encounter in daily life. The resulting assessment pool for the 1992 National Adult Literacy Survey included 82 document literacy tasks of which 26 were newly developed for the 1992 survey.

Previous sections of this chapter have described in detail and illustrated with examples the several criteria that must be taken into account when measuring the four variables associated with task difficulty on the document literacy scale. Just as with the 1990 survey and the 1985 young adult literacy assessment, these rules were applied to all document literacy tasks in the 1992 National Adult Literacy Survey. The resulting codes, along with RP80 task difficulties and IRT item parameters are shown in Table 13-14.

Table 13-14. List of new document literacy tasks, along with RP80 task difficulty, IRT item parameters, and values of variables associated with task difficulty: The 1992 National Adult Literacy Survey

Identifier	Task Description	Scaled RP80	IRT parameters			Complexity	Type of match	Distractor Plausibility	Information type
			a	b	c				
SCOR100	Social Security card: Sign name on line	70	0.505	-4.804	0.000	1	1	1	1
SCOR300	Driver's license: Locate expiration date	152	0.918	-2.525	0.000	2	1	2	1
N090301	Essence: Determine page certain article begins on	211	1.124	-1.225	0.000	3	1	2	2
N120601	Middle class: Find projected percent	213	0.795	-1.488	0.000	4	3	2	2
N010401	Vehicle chart: Find correct information	215	0.903	-1.341	0.000	2	1	2	2
N080802	Auto maintenance form: Enter given information	233	1.357	-0.684	0.000	3	2	3	2
N120101	Campus map: Mark map for given info	239	0.986	-0.802	0.000	7	3	2	1
N130103	S.S. card application: Identify and enter info(2)	243	2.106	-0.291	0.000	5	2	3	1
N130102	S.S. card application: Identify and enter info(3)	243	1.270	-0.544	0.000	5	2	2	3
N110302	Certified mail rec't: Enter postage and fee	244	0.714	-1.026	0.000	2	2	2	2
N110301	Certified mail rec't: Enter name and address	251	0.812	-0.743	0.000	2	2	2	2
N130104	S.S. card application: Identify and enter info(4)	251	2.159	-0.111	0.000	5	2	2	1
N090401	Essence: Determine topic of given article	257	0.988	-0.448	0.000	3	2	2	3
N130101	S.S. card application: Identify and enter info(1)	259	1.619	-0.096	0.000	5	2	2	2
N080701	Bus schedule: Mark map correctly for given info	260	1.095	-0.312	0.000	9	3	2	1
N010801	Trend chart: Mark information on chart	266	0.808	-0.463	0.000	3	3	2	1
N120201	Campus map: Find correct room for given dean	267	0.842	-0.403	0.000	7	3	2	2
N090501	Essence: Determine topic of section of magazine	285	0.671	-0.301	0.000	3	4	2	3
N100501	Opinions table: Mark sentence explaining action	304	1.039	0.486	0.000	4	3	2	4
N080601	Bus schedule: Take correct bus for given condition	305	1.040	0.505	0.000	9	4	2	2
N011001	Trend chart: Determine least # of points needed	317	0.646	0.261	0.000	3	5	3	2
N080801	Auto maintenance form: Enter information given (1)	323	0.763	0.570	0.000	3	3	4	2
N110701	Credit card table: Find correct bank	335	0.470	0.126	0.000	5	4	4	1
N100601	Opinions table: Find correct group for given info	343	1.135	1.285	0.000	4	2	4	2
N010901	Trend chart: Put information on chart	386	0.721	1.702	0.000	3	5	5	5
N110901	Credit card table: Give 2 differences	388	0.829	1.883	0.000	5	8	2	5
N100701	Summarize views of parents & teachers	396	1.128	2.300	0.000	4	8	3	5

Just as with the prose literacy tasks, parallel analysis of the document literacy tasks used in the 1992 National Adult Literacy Survey provides another piece of validation evidence. The four task variables were again used in a regression analysis designed to predict the RP80 scale values of each task. Table 13-15 shows the results of these analyses for both the new document literacy tasks as well as the entire 1992 pool of document literacy tasks. Overall, the three process variables were significant for both the new 1992 tasks and for the complete set of document literacy tasks. Just as on the prose scale, the same variables contribute to the predictive models, and the amount of variance accounted for is similar—92 percent on the 1990 survey and 88 percent on the 1992 survey.

Table 13-15. Unstandardized regression coefficients and standard errors predicting document literacy RP80 task difficulties on the basis of four structure and process variables: 25 new document literacy tasks and 81 total document literacy tasks from the 1992 survey of adults

Document literacy	New 1992 Tasks			All 1992 Tasks		
	Coeff	StdErr	p	Coeff	StdErr	p
Structure Variable						
Structural Complexity	2.83	2.58	.29	.09	1.07	.42
Process Variables						
Type of match	17.09	3.57	.00	20.33	2.27	.00
Plausibility of distractors	28.15	5.30	.00	22.19	2.74	.00
Type of information	9.84	5.07	.07	11.69	3.05	.00
Variance accounted for:						
R^2		88%			84%	
Adjusted R^2		85%			83%	
Degrees of freedom		20			76	

*Scoring was excluded from the regression equation due to the fact that it is an outlier.

Just as with the prose literacy scale, the equations predicting document literacy task difficulty in the 1992 data essentially reproduced the findings of the 1990 data. Again, there was no need to revisit the cutpoints for the document literacy levels, yet the language describing the literacy levels was improved in minor ways. Table 13-16 provides a comparison of the verbal descriptions of document literacy levels used in reporting the 1990 and 1992 survey results. The minor adjustments that can be seen in the 1992 descriptions were intended to clarify and systematize the language, but not to indicate any substantive changes in their meaning.

Table 13-16. Descriptions of document literacy levels based on type of match, plausibility of distractors, and type of information: 1990 survey of job-seekers and the 1992 National Adult Literacy Survey

Levels	1990 Description	1992 Description
Level 1 0-225	Tasks at this level are the least demanding. In general, they require the reader to either locate a piece of information based on a literal match or to enter information from personal knowledge.	Tasks in this level tend to require the reader either to locate a piece of information based on a literal match or to enter information from personal knowledge onto a document. Little, if any, distracting information is present.
Level 2 226-275	Tasks at this level begin to become more varied. Some still require the reader to match a single match a single piece of information; however, tasks occur where there are several distractors or where the match is based on low-level inferences. Tasks at this level also begin to require the reader to cycle through information or to integrate information.	Some tasks in this level require the reader to match a single piece of information; however, several distractors may be present, or the match may require low-level inferences. Tasks in this level may also ask the reader to cycle through information in a document or to integrate information from various parts of a document.
Level 3 276-325	Tasks at this level tend to require the reader to either integrate three pieces of information or to cycle through materials in rather complex tables or graphs in which distractor information is present.	Some tasks in this level require the reader to integrate multiple pieces of information from one or more documents, Others ask readers to cycle through rather complex tables or graphs which contain information that is irrelevant or inappropriate to the task.
Level 4 326-375	Tasks at this level tend to demand more from the reader. Not only are multiple-feature matching, cycling, and integration of information maintained, the degree of inferencing is increased. Cycling tasks often require the reader to make five or more responses with no designation of the correct number of responses. Conditional information is also present and must be taken into account.	Tasks in this level, like those at the previous levels, ask readers to perform multiple-feature matches, cycle through documents, and integrate information; however, they require a greater degree of inferencing. Many of these tasks require readers to provide numerous responses but do not designate how many responses are needed. Conditional information is also present in the document tasks at this level and must be taken into account by the reader.
Level 5 376-500	Tasks at this level require the most from the reader. The reader must search through complex displays contain[ing] multiple distractors, make high text-based inferences or use specialized knowledge.	Tasks in this level require the reader to search through complex displays that contain multiple distractors, to make high-level, text-based inferences, and to use specialized knowledge.

13.5.3 Quantitative Literacy

As with the prose and document literacy scale, the development of the 1992 National Adult Literacy Survey entailed both the reuse of existing quantitative literacy tasks from the 1985 young adult literacy assessment and the production of new quantitative literacy tasks. The new tasks continued the emphasis on prose texts and documents drawn from authentic sources that adults might ordinarily encounter in daily life. The resulting assessment pool for the 1992 National Adult Literacy Survey included 43 quantitative literacy tasks of which 28 were newly developed for the 1992 survey.

Previous sections of this chapter have described in detail and illustrated with examples the several criteria that must be taken into account when measuring the four variables associated with task difficulty on the quantitative literacy scale. Just as with the 1990 survey and the 1985 young adult literacy assessment, these rules were applied to all quantitative literacy tasks in the 1992 National Adult Literacy Survey. The resulting codes, along with RP80 task difficulties and IRT item parameters are shown in Table 13-17.

Table 13-17. List of quantitative literacy tasks, along with RP80 task difficulty, IRT item parameters, and values of variables associated with task difficulty: The 1992 National Adult Literacy Survey

Identifier	Quantitative Literacy Items	Scaled RP80	IRT parameters			Complexity	Type of match	Distractor Plausibility	Calculation type	Operation specificity
			a	b	c					
NC00501	Checking: Total bank deposit entry	192	0.661	-2.792	0.000	2	1	1	1	1
N110303	Certified mail rec't: postage and fees	239	0.790	-1.730	0.000	3	1	2	1	2
NC00601	Price diff: Sleuth & On the Town	247	0.718	-1.690	0.000	2	2	1	2	2
N100801	Salt River: Determine diff in costs	251	0.648	-1.738	0.000	5	2	2	2	3
N101001	Salt River: hours between points	278	0.944	-0.838	0.000	5	3	4	1	3
N090901	Carpet ad: Diff in reg and sale price	278	0.790	-1.004	0.000	2	3	1	2	4
N120701	Pct diff black & white middle class	280	0.909	-0.845	0.000	4	4	2	2	4
N090201	Get net total owed after deduction	284	1.677	-0.349	0.000	3	2	4	2	2
N010501	Vehicle chart: sum of percentages	287	0.851	-0.769	0.000	2	2	2	1	2
N110801	Credit card table: Difference in rates	300	0.882	-0.495	0.000	4	1	2	2	4
N130601	Rec room: num wall panels needed	307	1.112	-0.184	0.000	5	1	2	4	3
N090101	Discount if oil bill paid in 10 days	309	1.347	-0.018	0.000	3	2	2	3	5
N080501	Time: student union to 17th & Main	322	0.757	-0.248	0.000	9	4	2	2	5
N011101	Gas gauge: show calculations	330	1.035	0.196	0.000	2	3	1	5	6
N121001	Miles/day Butcher went	332	1.018	0.218	0.000	2	3	2	4	6
N100901	Salt River: Miles between stops	334	0.623	-0.264	0.000	5	2	4	1	5
N010701	Vehicle chart: magnitude of diff	342	1.034	0.411	0.000	2	3	4	4	3
N081001	Rank juices by expense, w/ reasons	344	0.733	0.122	0.000	2	4	1	3	6
N130901	Money needed to raise child	351	0.946	0.499	0.000	2	1	2	3	5
N010601	Vehicle chart: Solution to pct problm	355	1.122	0.717	0.000	2	1	2	3	6
N110201	Blood pressure: Death rate	360	1.033	0.741	0.000	2	1	2	3	6
N120801	Middle class: Diff in size of pct	366	1.013	0.831	0.000	4	3	2	4	4
N080401	Yrly amount for couple w/ basic SSI	369	0.696	0.521	0.000	2	2	4	3	6
N080901	Auto form: Calculate miles/gallon	376	0.851	0.856	0.000	3	3	2	5	4
N130501	RecRoom: Feet of molding needed	389	0.655	0.819	0.000	5	2	3	5	7
N121101	Butcher: diff in completion times	406	0.960	1.518	0.000	2	2	1	5	5
N091001	Carpet ad:Total cost to carpet room	421	0.635	1.371	0.000	2	1	2	5	7
N130701	Rec room: Describe solution	436	0.846	1.962	0.000	5	2	2	5	7

*Structural Complexity

As with the prose and document literacy tasks, parallel analysis of the quantitative literacy tasks used in the 1992 National Adult Literacy Survey provides more evidence of construct validity. The four task variables were again used in a regression analysis designed to predict the quantitative literacy RP80 scale values of each task. Table 13-18 shows the results of these analyses for both the new quantitative literacy tasks as well as the entire 1992 pool of quantitative literacy tasks. Overall, the task variables were significant for both the new 1992 tasks and for the complete set of quantitative literacy tasks. Just as on the prose and document literacy scales, the same variables contribute to the predictive models, and the amount of variance accounted for is similar—84 percent on the 1990 survey (Table 13-7) and 84 percent on the 1992 survey (Table 13-18). Moreover, the similarity of results between the pools of the 1990 and 1992 tasks is especially important as evidence in support of the appropriateness of interpreting the 1992 survey results in terms of the five levels of literacy proficiencies established for reporting the 1990 survey results.

Table 13-18. Unstandardized regression coefficients and standard errors predicting quantitative literacy RP80 task difficulties on the basis of five structure, process, and formulate variables: 28 new quantitative literacy tasks and 43 total quantitative literacy tasks from the 1992 survey of adults

Quantitative literacy	New 1992 Tasks			All 1992 Tasks		
	Coeff	StdErr	p	Coeff	StdErr	p
Structure Variable						
Structural Complexity	.33	3.18	.92	.78	2.81	.78
Process Variables						
Type of match	-2.57	5.16	.62	.68	4.24	.87
Plausibility of distractors	12.25	5.16	.03	9.66	4.17	.03
Formulate Variables						
Type of calculation	20.75	4.62	.00	14.43	3.21	.00
Operation specificity	16.64	3.84	.00	18.39	2.94	.00
Variance accounted for:						
R^2		84%			81%	
Adjusted R^2		81%			78%	
Degrees of freedom		22			37	

As with the prose and document literacy scales, the equations predicting quantitative literacy task difficulty in the 1992 data essentially reproduced the findings of the 1990 data. Again, there was no need to revisit the cutpoints for the quantitative literacy levels, yet the language describing the literacy levels was adjusted in minor ways without affecting the substantive meaning of the descriptions. Table 13-19 provides a comparison of the verbal descriptions of quantitative literacy levels used in reporting the 1990 and 1992 survey results.

Table 13-19. Descriptions of quantitative literacy levels based on type of operation, plausibility of distractors, and type of calculation: 1990 survey of job-seekers and the 1992 National Adult Literacy Survey

Levels	1990 Description	1992 Description
Level 1 0-225	Although no quantitative tasks used in this assessment fall within this level, experience suggests that such tasks would require a single, relatively simple operation for which the numbers are given and the operation specified.	Tasks in this level require readers to perform single, relatively simple arithmetic operations, such as addition. The numbers to be used are provided and the arithmetic operation to be performed is specified.
Level 2 226-275	Tasks at this level typically require the use of a single operation based on numbers that are either stated in the question or easily located in the material. In addition, the operation needed is either stated in the question or easily determined based on the format of the problem—for example, entries on a bank deposit slip or order form.	Tasks in this level typically require readers to perform a single operation using numbers that are either stated in the task or easily located in the material. The operation to be performed may be stated in the question or easily determined from the format of the material (for example, an order form).
Level 3 276-325	What appears to distinguish tasks at this level is that two or more numbers needed to solve the problem must be found in the stimulus material. Also the operation(s) needed can be determined from arithmetic relation terms.	In tasks in this level, two or more numbers are typically needed to solve the problem, and these must be found in the material. The operation(s) needed can be determined from the arithmetic relation terms used in the question or directive.
Level 4 326-375	Quantitative tasks at level 4 tend to require two or more sequential operations or the application of a single operation where either the quantities must be located in complex displays and/or the operation must be inferred from semantic information given or prior knowledge.	These tasks tend to require readers to perform two or more sequential operations or a single operation in which the quantities are found in different types of displays, or the operations must be inferred from semantic information given or drawn from prior knowledge.
Level 5 376-500	Quantitative tasks at this level are the most demanding. They tend to require the reader to perform multiple operations and to disembed features of a problem from stimulus material or to rely on background knowledge to determine the quantities or operations needed.	These tasks require readers to perform multiple operations sequentially. They must disembed the features of the problem from text or rely on background knowledge to determine the quantities or operations needed.

13.5.4 Levels of Prose, Document, and Quantitative Literacy

In addition to the above analyses of continuous RP80 task difficulties, it was also important to analyze the 1992 survey tasks in terms of categorical literacy levels to make sure that the variables and amounts of explained variance were roughly the same. The results of these regression analyses are shown in table 13-20; the R^2s are in the range from 79 to 88 percent. These results are quite comparable with those reported in the 1985 and 1990 surveys. In addition, the amount of variance accounted for is nearly identical for the level regressions compared to the RP80 difficulty regressions—89 percent for continuous tasks compared to 88 percent for levels on the prose scale; 88 percent compared to 86 percent respectively on the document scale; and 81 percent compared to 79 percent on the quantitative scale.

Table 13-20. Unstandardized regression coefficients and standard errors predicting five levels of RP80 task difficulty on the basis of six structure and process variables: All prose, document, and quantitative literacy tasks from the 1992 survey of adults

	Prose literacy			Document literacy			Quantitative literacy		
	Coeff	StdErr	p	Coeff	StdErr	p	Coeff	StdErr	p
Structure Variable									
Structural Complexity	.01	.03	.65	.00	.20	.93	.07	.37	.07
Process Variables									
Type of match	.53	.06	.00	.42	.04	.00	.04	.08	.64
Plausibility of distractors	.26	.06	.00	.43	.05	.00	.21	.79	.01
Type of information	.15	.07	.05	.23	.05	.00	-	-	-
Formulate Variables									
Type of calculation	-	-	-	-	-	-	.21	.06	.00
Operation specificity	-	-	-	-	-	-	.35	.06	.00
Variance accounted for:									
R^2		88%			86%			79%	
Adjusted R^2		86%			86%			76%	

13.6 CONCLUSION

One of the goals of large-scale surveys is to provide a set of information that can inform the decision-making process. Important to this goal is presenting data in a manner that will enhance the understanding of what has been measured and of the conclusions that may be drawn both within and across assessments. The theoretical model that has evolved through three literacy assessments using the same definition and measurement framework has been a useful and valid way to report on the condition of adult literacy in America. This model identifies a set of variables that has been shown to underlie successful performance on a broad array of literacy tasks across several surveys. These variables, in turn, have been useful in developing new literacy tasks that help us to refine and extend our measurement of literacy. Moreover, they provide a framework for understanding what is being measured that allows us to identify levels of performance that have generalizability and validity across assessments and groups, rather than interpreting results in terms of discrete tasks.

Collectively, the knowledge and understanding that derives from such models contributes to an evolving conception of test design that begins to move away from merely assigning a numerical value (or position) to an individual based on responses to a set of tasks and toward assigning meaning and interpretability to this number. This, in turn, provides evidence of the appropriateness of the theoretical models predicted to underlie consistencies in performance.