DOCUMENT RESUME

ED 222 577 TM 820 754

AUTHOR	Croll, Paul R.
TITLE	Computerized Adaptive Testing System Design: Preliminary Design Considerations.
INSTITUTION	Office of Personnel Management, Washington, D.C.
SPONS AGENCY	Navy Personnel Research and Development Center, San Diego, Calif.
REPORT NO	NPRDC-TR-82-52
PUB DATE	Jul 82
NOTE	53p.; Available in microfiche only, due to small print in figures and appendix.
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Adults; *Computer Assisted Testing; *Computer Programs; Microcomputers; Minicomputers; *Occupational Tests; *Psychological Testing; *Systems Development
IDENTIFIERS	*Adaptive Testing; Department of Defense

ABSTRACT

A functional design model for a computerized adaptive testing (CAT) system was developed and presented through a series of hierarchy plus input-process-output (HIPO) diagrams. System functions were translated into system structure: specifically, into 34 software components. Implementation of the design in a physical system was addressed through brief discussions of hardware, software, interfaces, and personnel requirements. Further steps in CAT system development were identified, including design testing, evaluation, and refinement. Both micro- and mini-computer-based hardware configurations were evaluated and found capable of supporting test administration and station monitoring. The functional design model and the system structure specified in this report were recommended for the Department of Defense CAT system. (Author)

NPRDC TR 82-52 July 1982

COMPUTERIZED ADAPTIVE TESTING SYSTEM DESIGN: PRELIMINARY DESIGN CONSIDERATIONS

Paul R. Croll

Office of Personnel Management
Washington, DC 20415

Reviewed by
Martin F. Wiskoff

Released by
James F. Kelly, Jr.
Commanding Officer

Navy Personnel Research and Development Center
San Diego, California 92152

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPRDC TR 82-52	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) COMPUTERIZED ADAPTIVE TESTING SYSTEM DESIGN: PRELIMINARY DESIGN CONSIDERATIONS		5. TYPE OF REPORT & PERIOD COVERED Interim Report 1 Oct 1979-30 Sep 1981
		6. PERFORMING ORG. REPORT NUMBER 12-81-1
7. AUTHOR(s) Paul R. Croll		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Personnel Research and Development Center Office of Personnel Management Washington, DC 20415		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS I - EU
11. CONTROLLING OFFICE NAME AND ADDRESS Navy Personnel Research and Development Center San Diego, California 92152		12. REPORT DATE July 1982
		13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Computerized adaptive testing
Tailored testing
Computer system design
Psychological testing
Personnel testing

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A functional design model for a computerized adaptive testing (CAT) system was developed and presented through a series of hierarchy plus input-process-output (HIPO) diagrams. System functions were translated into system structure; specifically, into 34 software components. Implementation of the design in a physical system was addressed through brief discussions of hardware, software, interfaces, and personnel requirements. Further steps in CAT system development were identified, including design testing, evaluation, and refinement. Both micro- and minicomputer-based hardware configurations were evaluated and found capable of supporting test administration and station monitoring. The functional design model and the system structure specified in this report were recommended for the Department of Defense CAT system.

FOREWORD

A joint-service coordinated effort is in progress to develop a computerized adaptive testing (CAT) system and to evaluate its potential for use in the Military Enlistment Processing Stations as a replacement for the Armed Services Vocational Aptitude Battery (ASVAB) printed tests. The Navy Personnel Research and Development Center has been designated lead laboratory for this effort.

This report describes the preliminary design considerations that were incorporated into the government's formal solicitation of proposals for CAT system design and development. A previous report (NPRDC Tech. Note 82-22) described the functional requirements and objectives of the CAT system.

The contracting officer's technical representative was Dr. James R. McBride.

JAMES F. KELLY, JR.
Commanding Officer

JAMES J. REGAN
Technical Director

SUMMARY

Problem

Much research has been conducted, both within and outside the Department of Defense (DoD), on the psychometric underpinnings of computerized adaptive testing (CAT). In January 1979, a DoD joint-service effort was initiated to evaluate the feasibility of implementing a CAT system for enlisted personnel accession testing. As the lead laboratory directing the effort, NAVPERSRANDCEN has primary responsibility for the design, development, testing, and evaluation of such a CAT system.

Objectives

The objectives of this effort were to:

1. Establish the principles on which the tailored testing system will be developed.

2. Develop a functional design model for the CAT system, including specification of its functional components and their structural relationships, as well as design implications for the physical system.

Approach

A top-down structural design technique called hierarchy plus input-process-output (HIPO) was used in developing the CAT system functional design model. Functional requirements specified by NAVPERSRANDCEN, as well as experience gained in the design of a similar system for the Office of Personnel Management, were used to delineate the functions that should be performed by the system and the way in which those functions should interface. The current technical literature on computer hardware was reviewed to assess implications of the functional design for the physical system. A loosely coupled microprocessor configuration was compared with shared minicomputer configurations for single-site hardware support.

Results

1. Application of the HIPO approach to the design of the CAT system resulted in the initial design level specification of four major functional subsystems comprised of 25 subfunctions of varying levels of specificity. The four major subsystems are (a) item banking, (b) measurement control, (c) test administration and scoring, and (d) monitoring and quality control.

2. Thirty-four software components were specified by system function.

3. Internal and external system interfaces were identified, detailing data and control paths among the four major functional subsystems and the Military Enlistment Processing Station Reporting System.

4. Personnel considerations for system operation were specified, describing the desired minimum system impact on both operating personnel and examinees.

5. Further steps in CAT system development were identified, including the need for testing, evaluation, and refinement of the system design as part of the continuing process of system development.

7

6. A review of the state of the art in computer hardware and a comparison of microprocessors and minicomputers showed that both were capable of supporting CAT interactive testing and monitoring functions.

Recommendations

1. The CAT system design should be based on the 4 major functional subsystems and 25 subfunctions specified in this report.

2. The HIPO approach should continue to be employed throughout the evolution of the final system design.

3. Both microprocessors and minicomputers should be evaluated for support of CAT test administration and for station-monitoring functions.

4. The 34 software components identified in this report should serve as the basis for system software development.

5. FORTRAN, Pascal, or another high-level structured programming language should be chosen for software development.

6. Personnel requirements for system operation should be minimized.

7. Procedures for design testing, evaluation, and refinement should be specified and implemented in the CAT system development process.

8

CONTENTS

Page

LIST OF FIGURES

INTRODUCTION

Background and Problem

The military services have, over many years, pursued innovative solutions to pressing personnel measurement problems. Since 1917, when the need for rapid classification of recruits resulted in the development of the first group intelligence tests, the military services have provided a major impetus to the development of new measurement technology (Anastasi, 1976). The huge selection and classification task brought on by World War II led to the development of the first multiple-ability aptitude batteries and brought recognition of the need for continuing research and development in selection and classification. The use of group tests, however, has meant some sacrifice of the accuracy provided by individualized tests. Recent research has sought to provide the measurement advantages of an individualized testing procedure (in the mold of the early Binet tests), while retaining the administrative efficiencies associated with group tests. Computerized adaptive testing (CAT) is the outgrowth of that research.

CAT is a remarkably effective combination of recent developments in latent trait theory and of continuing advances in computer technology (Urry, 1977a). Unlike conventional paper-and-pencil group testing, in which identical test forms are administered simultaneously to large groups of examinees, CAT is an individualized testing procedure that constructs, administers, and scores tests interactively during the testing session. In conventional group testing, enough test questions must be included to assess all levels of ability in the population of applicants. As a result, examinees must answer many questions that are inappropriate to their own levels of ability. In CAT, examinees receive only those questions appropriate to their own levels of ability. The result is a test that is "adapted" or "tailored" to each examinee's level. Considerably fewer questions are required in CAT than in the group test to produce an estimate of ability at the same level of reliability.

The adaptive nature of the CAT procedure may be illustrated by the following scenario: The examinee sits at a testing station that consists of a video display and a keyboard and that may communicate with a remote computer or contain a dedicated microcomputer. When a test question appears on the video display screen, the examinee indicates an answer by pressing the appropriate key on the keyboard. If the answer is correct, a more difficult question is presented. If the answer is incorrect, an easier question is presented. With each succeeding response, the computer makes a revised estimate of the examinee's ability. As the testing sequence proceeds, each estimate becomes more reliable. The test is terminated when a previously specified level of reliability is reached. The procedure for multiple-ability testing is similar. This scenario would be repeated for each ability to be tested.

The apparent simplicity of this procedure belies the extreme complexity of its psychometric underpinnings (see Urry, 1981a, b). This complexity, coupled with the need for great accuracy in the accession testing process, presents the system-design challenge in CAT system development.

Exploratory and advanced development of CAT applications has been conducted at the Civil Service Commission (now the Office of Personnel Management (OPM)) (Clark, 1976; Urry, 1977a) and, more recently, at the Educational Testing Service (Lord, 1977a, b) the Air Force Human Relations Laboratory (Ree & Jensen, 1980), the Army Research

Institute (McBride, 1979), NAVPERSRANDCEN (McBride, 1980), and several universities.[1] In January 1979, the Department of Defense (DoD) established a joint-service project to develop a CAT system and evaluate its potential for use in the Military Enlistment Processing Stations (MEPS) (formerly the Armed Forces Examining and Entrance Stations (AFEES)) as a replacement for the Armed Services Vocational Aptitude Battery (ASVAB), which is used for enlisted personnel accession testing. As lead laboratory in this effort, NAVPERSRANDCEN has primary responsibility for design, development, testing, and evaluation of the CAT system.

The joint-service project has been conceived as a large-scale system development effort, integrating psychometric and engineering developments to meet system goals. This report is the second of a series that will result from the project. The first (McBride, 1982) described the functional requirements and objectives of the CAT system.

Objectives

The objectives of the effort reported here were to:

1. Establish the principles on which the tailored testing system will be developed.

2. Develop a functional design model for the CAT system, including specification of its functional components and their structural interrelationships, as well as design implications for the physical system.

APPROACH

Development of CAT System Functional Design Model

System Design Principles

The primary objectives of the CAT system development effort are the design, development, testing, and evaluation of a system for automated adaptive administration of DoD enlisted personnel selection and classification tests. The desired outcome of the development effort is an integrated set of well-defined inputs, processes, and outputs that meet the following criteria:

1. User (i.e., military service) needs may be easily translated into specifications that both define system products and provide control of system processes.

2. System products completely and consistently conform to user specifications.

3. System processes and products are continuously monitored to ensure such conformance.

The capability for delivery of well-defined products, meeting user needs and monitored for conformance with user specifications, is the essence of the CAT system.

---

[1] Several conferences have included work in this area. See Holtzman (1970), Clark, (1976), and Weiss (1978, 1980).

The system development problem has been approached through two distinct lines: (1) psychometric development of the procedures for adaptive testing and (2) engineering development of the physical system through which these procedures will be implemented. The application of system design principles to the development of the computer-based physical system is straightforward and well supported by present practice. The application of such principles to the development of psychometric procedures is unique, however, and can present a subtle danger to the integrity of the system as a whole.

The danger lies in the possible failure to recognize that the CAT system must be designed to meet psychometric objectives first. Engineering objectives must not be permitted to drive the system development effort. For example, modification of well proven CAT algorithms, based solely on an initial conception of hardware performance characteristics, is inappropriate. Rather, algorithmic requirements should, within reason, dictate hardware specifications. Viewing CAT system development as simply another data-processing system exercise is likely to compromise its psychometric integrity. Recognition of the tremendously complex network of interaction underlying systems design is especially necessary for CAT. System designers must understand the relationships among the system's psychometric and physical components. Appreciation of these relationships is critical to integrating the components into a properly functioning system.

To facilitate such integration, the design strategy chosen for the CAT system has focused on function rather than structure. Katzan (1976) describes a system function as a process that accepts one or more inputs and produces one or more outputs. The application of this definition in computer hardware or software design is straightforward. For example, the "multiply" function of a central processing unit (CPU) chip accepts a multiplier and a multiplicand, each of fixed length, and returns a product. Valid input sources and output destinations are inherent in the chip design. The application in software design is analogous, with the program code determining input sources and characteristics, output destinations and characteristics, and the intervening processing steps necessary to produce output from input. The application of this definition to the design of a psychometric system is less obvious. Even Chapanis (1970a, b), writing about human factors in systems engineering in de Greene's Systems Psychology, neglects to apply system design principles in developing psychometric procedures. Systems thinking is applied only to the problem of personnel selection and classification and then only in the sense that a systematic approach to selecting, evaluating, and training personnel is seen as a component of a larger design. Systems thinking need not stop short with the human factors or engineering psychology approach, however. It is readily applicable to basic psychometric developments as well.

If one defines a personnel measurement procedure as the administration, scoring, and evaluation of the results of a test of some ability, questions couched in system design terms can easily be raised. What are the desired outputs? Test records, scores, selection decisions? What are the processes required to obtain those outputs? Administering test questions, recording examinee responses, scoring, applying selection rules? What are the inputs required by the specified processes to produce the desired outputs? Instruction sets, test questions, examinee responses, scoring keys? This simplistic example illustrates the principle that psychometric issues such as personnel measurement may be addressed from a system design perspective, bringing to bear all the tools and techniques of that discipline. The design of a CAT system is a far more complex undertaking, but the development of a functional design model for the system greatly simplifies the dual tasks of psychometric and engineering development and facilitates their eventual integration.

For this effort, a functional design model was developed to address both the psychometric and the administrative or operational requirements of CAT and presented through a series of hierarchy plus input-process-output (HIPO) diagrams (IBM, 1975; Katzan, 1976).[2] The HIPO package consists of (1) a visual table of contents, (2) overview diagrams, and (3) detail diagrams. These components are described below and illustrated in the following section.

1. Visual Table of Contents. This snapshot of the system is a hierarchy diagram that presents a structured decomposition of system functions into subfunctions of increasing detail as the diagram is read from top to bottom. Reading from left to right across any level in the hierarchy diagram provides a description of what the system does at that level of detail. Also, outputs of a functional component generally serve as inputs to the component on its immediate right. The boxes in the hierarchy diagram contain the names and identification numbers of the overview and detail diagrams in the HIPO package. To obtain the description of a specific function or subfunction, the reader goes to the overview or detail diagram referenced in the visual table of contents.

2. Overview Diagrams. Overview diagrams are the most general descriptions of system function contained in the HIPO package. They take the form of input-process-output diagrams, with the inputs listed in the left block, the process steps in the middle block, and the outputs in the right block. These general diagrams merely list inputs, outputs, and steps; they provide no indication of how the inputs and outputs are related to the process steps, nor do they specify the precise form of the input and outputs. When steps in the process block are boxed, with identification numbers appearing in the lower right-hand corner of the box, they represent subfunctions and refer to lower level overview or detail diagrams describing the function.

3. Detail Diagrams. Detail diagrams describe system function more specifically than overview diagrams. They, too, take the form of input-process-output diagrams and generally describe system subfunctions. Inputs and outputs are described in more detail than in overview diagrams and are linked with the steps in the process block in which they are used. References to lower level subfunctions are similar to those in overview diagrams. Additionally, when the process being described will be implemented primarily in software, steps in the process block may point to internal and external subroutines.

System Design Stages

Several stages normally constitute any system development effort. These stages, which, collectively, are often called the system life cycle, include (modified from de Greene, 1970; Rubin, 1970): (1) problem definition, (2) requirements analysis, (3) concept development, (4) preliminary system design, (5) design testing, evaluation, and refinement, (6) system development, (7) system installation, (8) system operation, and (9) system modification or replacement. These stages are described in the following paragraphs.

1. Problem Definition. Problem definition, which provides the rationale either for modifying what already exists or for creating something new, must precede the development of any system. In the CAT system development effort, the problem has been defined as the elimination or amelioration of several problems and deficiencies inherent in

---

[2] The development of a functional design model for a CAT system has been based on analysis of the requirements specified by NAVPERSRANDCEN, as well as the author's experience with design of a similar system at OPM (see Croll & Urry, 1975).

the present paper-and-pencil versions of ASVAB (McBride, 1982). These problems include: (a) excessive duration of personnel test sessions, (b) poor measurement precision at high and low ability levels, (c) susceptibility to theft, compromise, and coaching, (d) expense of printing, storage, and distribution for multiple forms of test booklets and answer sheets, (e) susceptibility to errors inherent in manual score tallying, score conversion, computation of score composites, and score recording, and (f) long lead time and high expense needed to develop replacement forms. The apparent capability of CAT technology to provide a single solution to these problems led to its selection as the technology of choice in developing a replacement for the present ASVAB.

2. Requirements Analysis. Requirements analysis provides clear definition of system objectives and serves as the basis for specifying system functions. System requirements can be many and varied. Categories of CAT system requirements include psychometric, administrative and operational, physical system performance, reliability, security, maintenance, personnel, training, documentation, and interface requirements. The definition of system requirements not only serves as the basis for system design but also allows system evaluation criteria to be specified.

3. Concept Development. A description of the system, a rough approximation, is produced in the concept development stage. Several preliminary design concepts may be proposed and evaluated, resulting in selection of a single candidate concept. Concept development bridges the specification of system objectives and the development of detailed design specifications. It allows one to think through design considerations before making a commitment to a specific system design. Descriptions of operational scenarios, functions of system elements, physical system configurations, system interfaces, and personnel considerations are usually provided as part of the system's design concept.

4. Preliminary System Design. The system design concept is refined into a set of hierarchical functional descriptions of system components and their interrelationships. Those detailed descriptions serve as the basis for design of the system's structure, its prototyping, and its final system development. As indicated previously, such functional descriptions were developed using the HIPO technique, which describes system functions in terms of inputs, processes, and outputs. These descriptions are presented hierarchically, showing in progressively greater detail the functional relationships among system components. All required inputs, processes, and outputs at each level of functional detail are specified.

5. Design Testing, Evaluation, and Refinement. Once the preliminary system design is completed, it must be tested, evaluated, and refined. A working model of the system, based on the preliminary design, is constructed and then tested and evaluated to validate the design against systems objectives. This prototype should be an accurate representation of what the system will look like and how it will perform when it is placed into operation. The prototype must be carefully evaluated, taking care to ensure that evaluation criteria have been well specified and that the test and evaluation process accurately simulates real-world conditions. This stage further allows design refinement, so that deviations from system objectives or evaluation criteria may be corrected before full-scale system development begins.

6. System Development. Full-scale implementation of the system design includes the final development of all system components, interfaces, operating procedures, personnel requirements, and system documentation. This stage focuses primarily on the physical system and its support requirements and is the final embodiment of the functional design. At the completion of this stage, the system is ready for installation in the operating environment.

7. System Installation. When the system is placed in the operating environment, it is not unusual for the system design to be validated further through operational field testing and evaluation. When the system has been validated in the actual operating environment, it may be fully deployed for operation. This stage also includes completion of training requirements for all system personnel.

8. System Operation. After installation and deployment, the ongoing stage of system operation includes not only day-to-day operation but also monitoring and quality control. In CAT system operation, it would also include periodic updating of the question files (item bank) from which test questions are selected, as well as selected presentation of experimental test questions for research purposes.

9. System Modification or Replacement. Any system has a finite life. Changing requirements, new technology, or system evolution may dictate modifications or replacement. The key issue in this stage is awareness of change coupled with careful planning, so that required changes may proceed smoothly.

These stages in the system life cycle provide the perspective for discussion of preliminary design considerations. The first five stages provide the essential principles upon which a good system design will be based. The use of the HIPO technique simplifies the task of integrating psychometric and engineering developments into an efficient CAT system.

Literature Review

The current technical literature on computer hardware was reviewed to assess implications of the functional design for the physical system.

RESULTS

CAT System Functions

In CAT, tests are constructed, administered, and scored interactively during the testing session. What functions are necessary to this process? First, it is obvious that a function encompassing test construction, administration, and scoring is needed. Test questions for each ability are selected from an item bank. Item banks are carefully constructed sets of test questions having well specified psychometric properties; each item bank is designed to measure a single ability. Thus, a function providing for item banking must also be defined. In CAT, a test may be terminated when a specified level of reliability is reached. Because multiple-ability testing may require a weighted composite score, a function providing termination rules and score weights is necessary. A function that monitors CAT functioning and quality control reporting is needed to let the user know when things go wrong.

By applying such a simple functional analysis to the CAT process, four major functions were identified: (1) item banking, (2) measurement control, (3) test administration and scoring, and (4) monitoring and quality control. These functions were formally expressed using the HIPO technique. The visual overview of the CAT system is provided in Figure 1; and the system overview diagram, in Figure 2. Outputs of the item banking and the measurement control components are required as inputs to the test administration component, and outputs from the test administration component are required as inputs for monitoring and quality control. These functions and their associated subfunctions are further specified in the detail diagrams for the functions (Figures 3 through 17) and are described commencing on page 20.

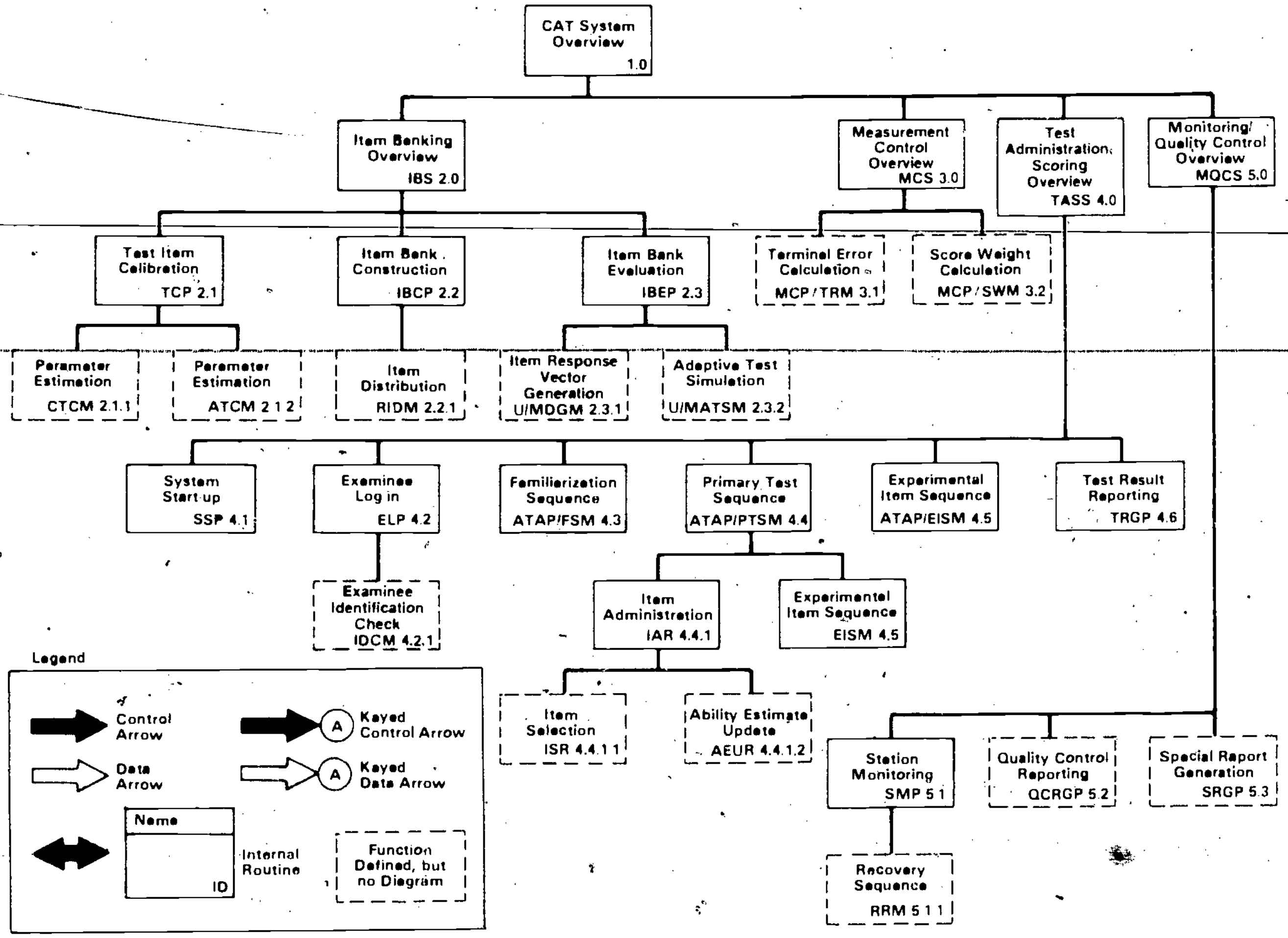

Figure 1. Visual table of contents for the DoD CAT system's initial design level.

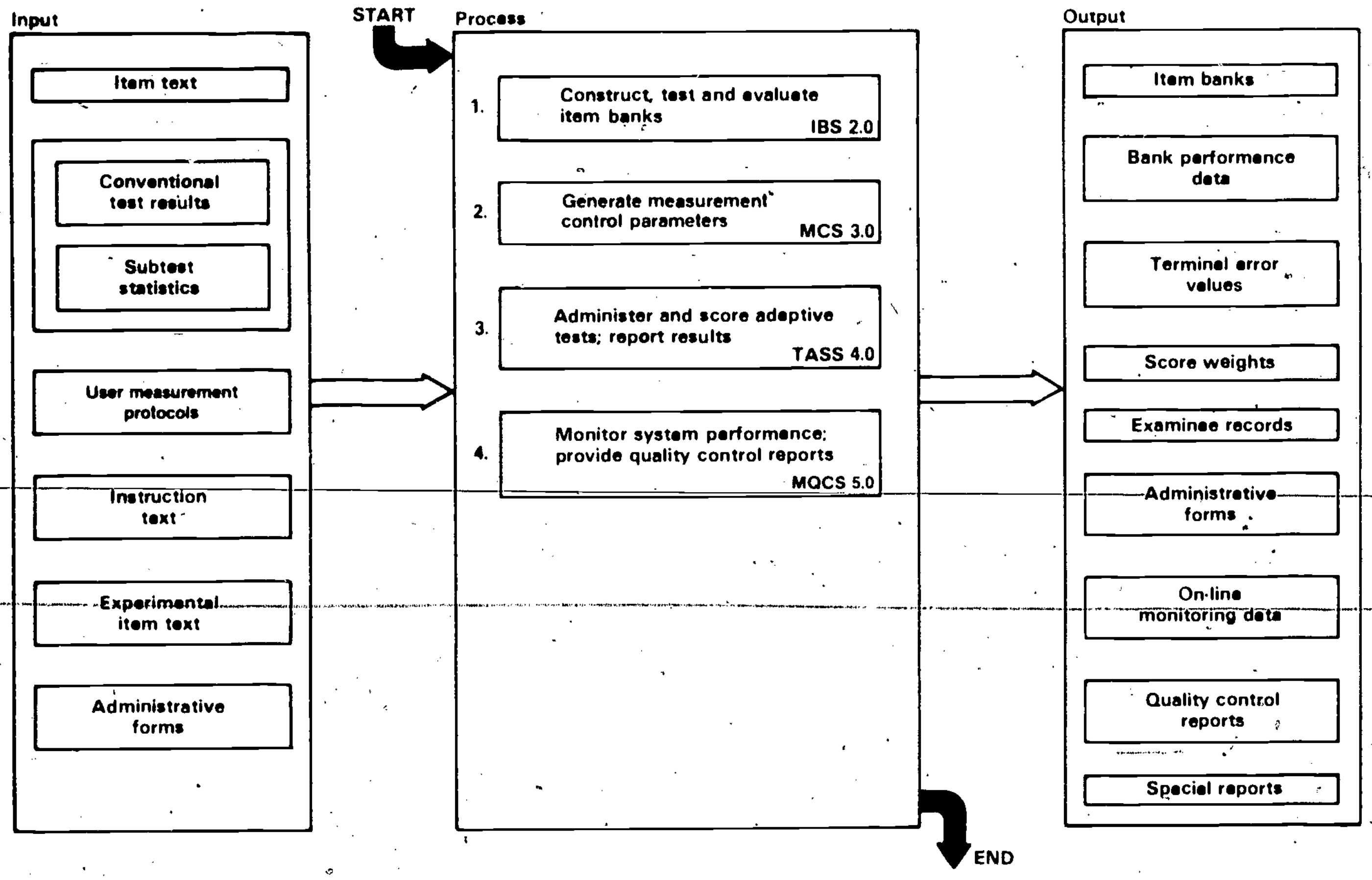

Figure 2. Functional overview of the DoD CAT system.

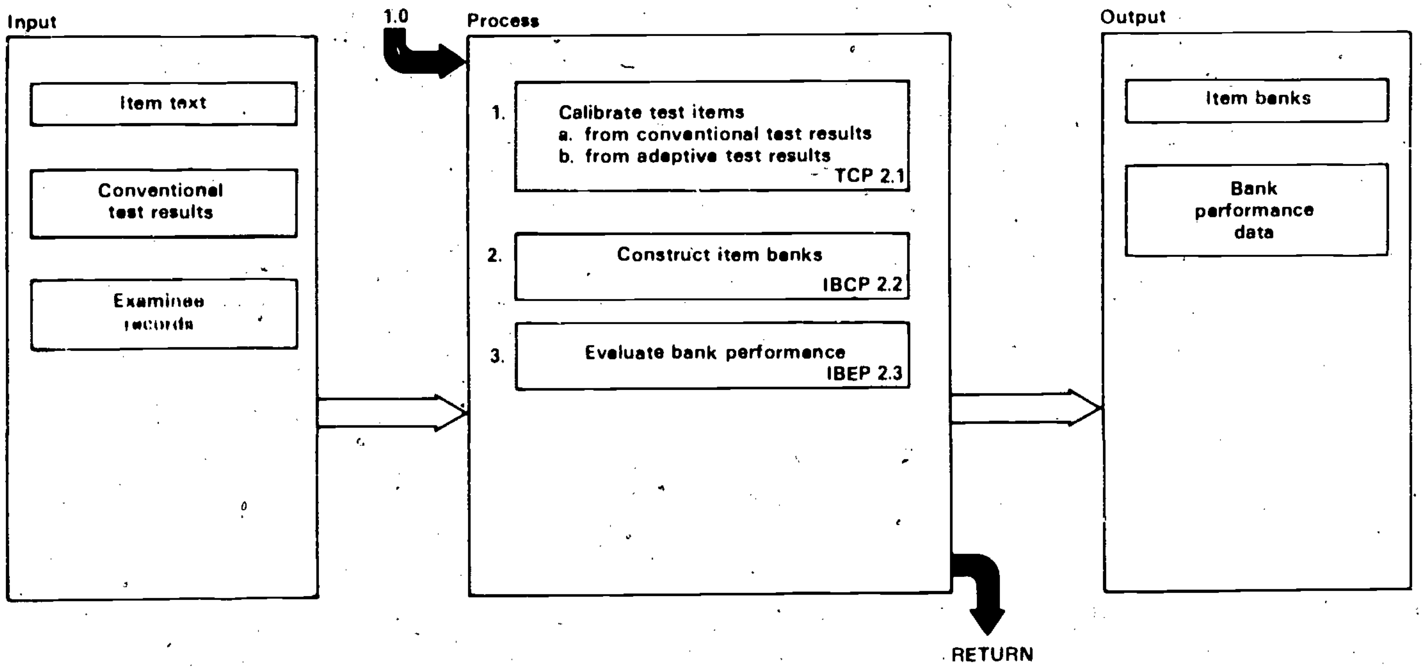

Figure 3. The item banking function.

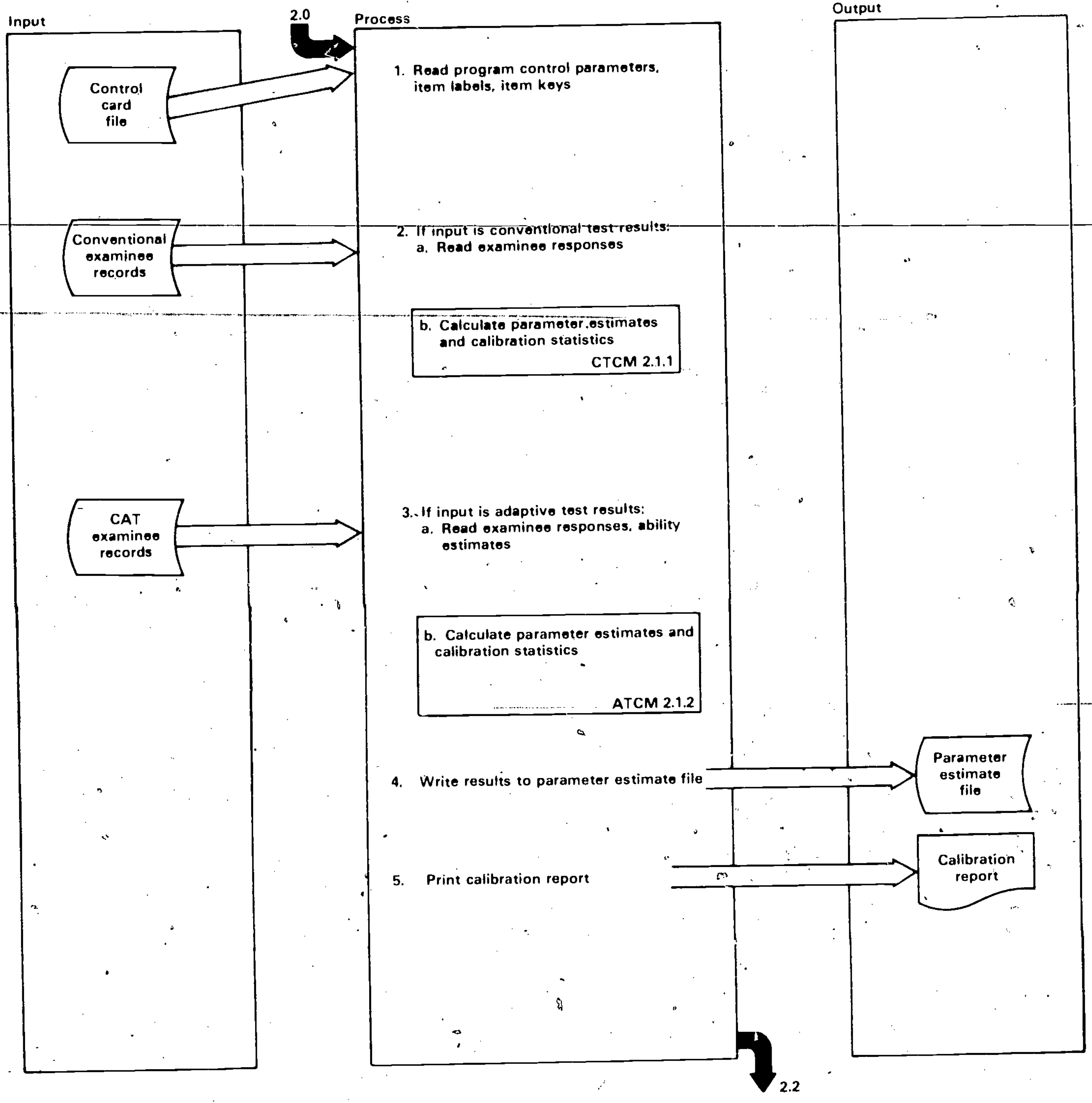

Figure 4. The test item calibration subfunction.

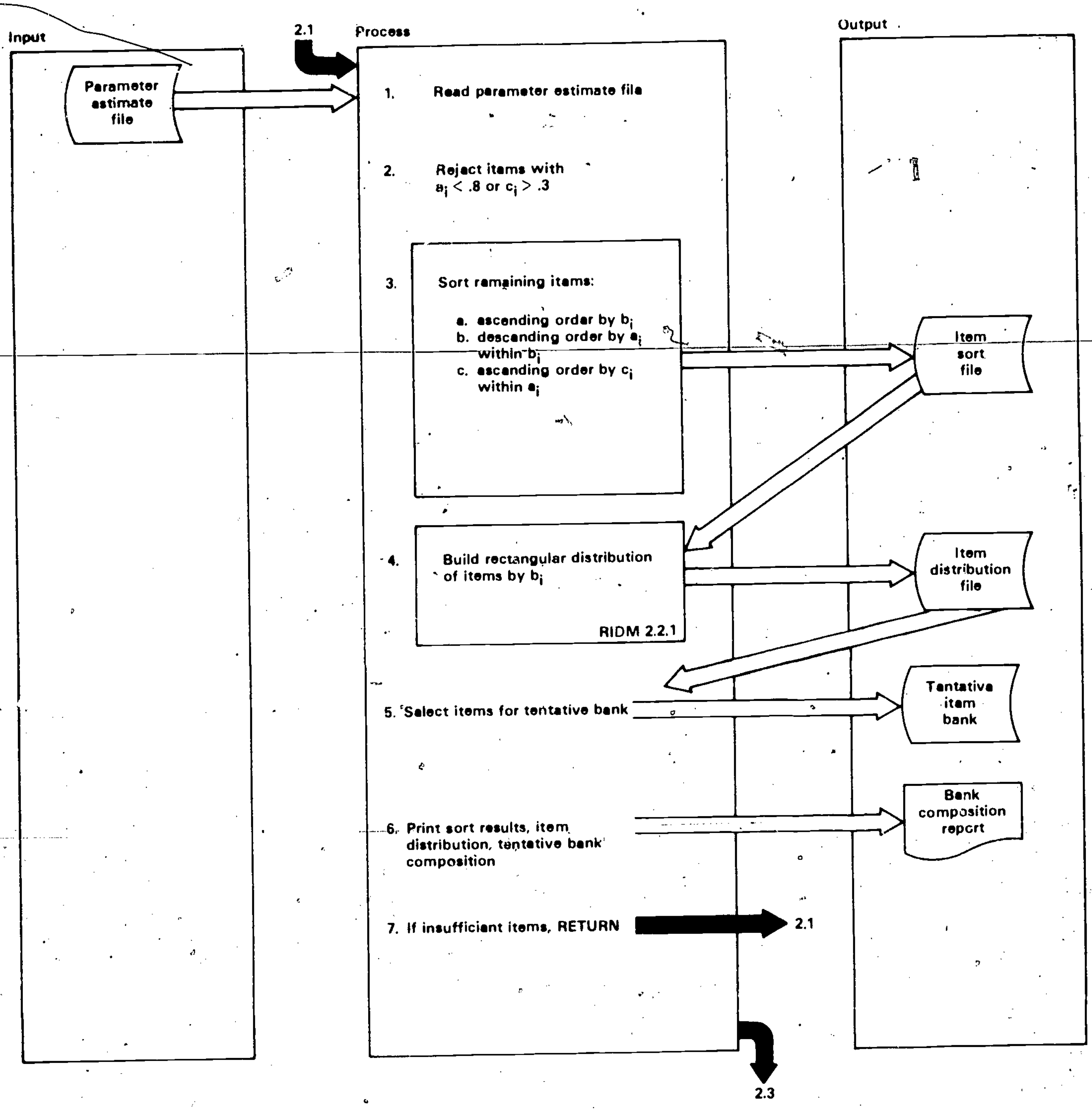

Figure 5. The item bank construction subfunction.

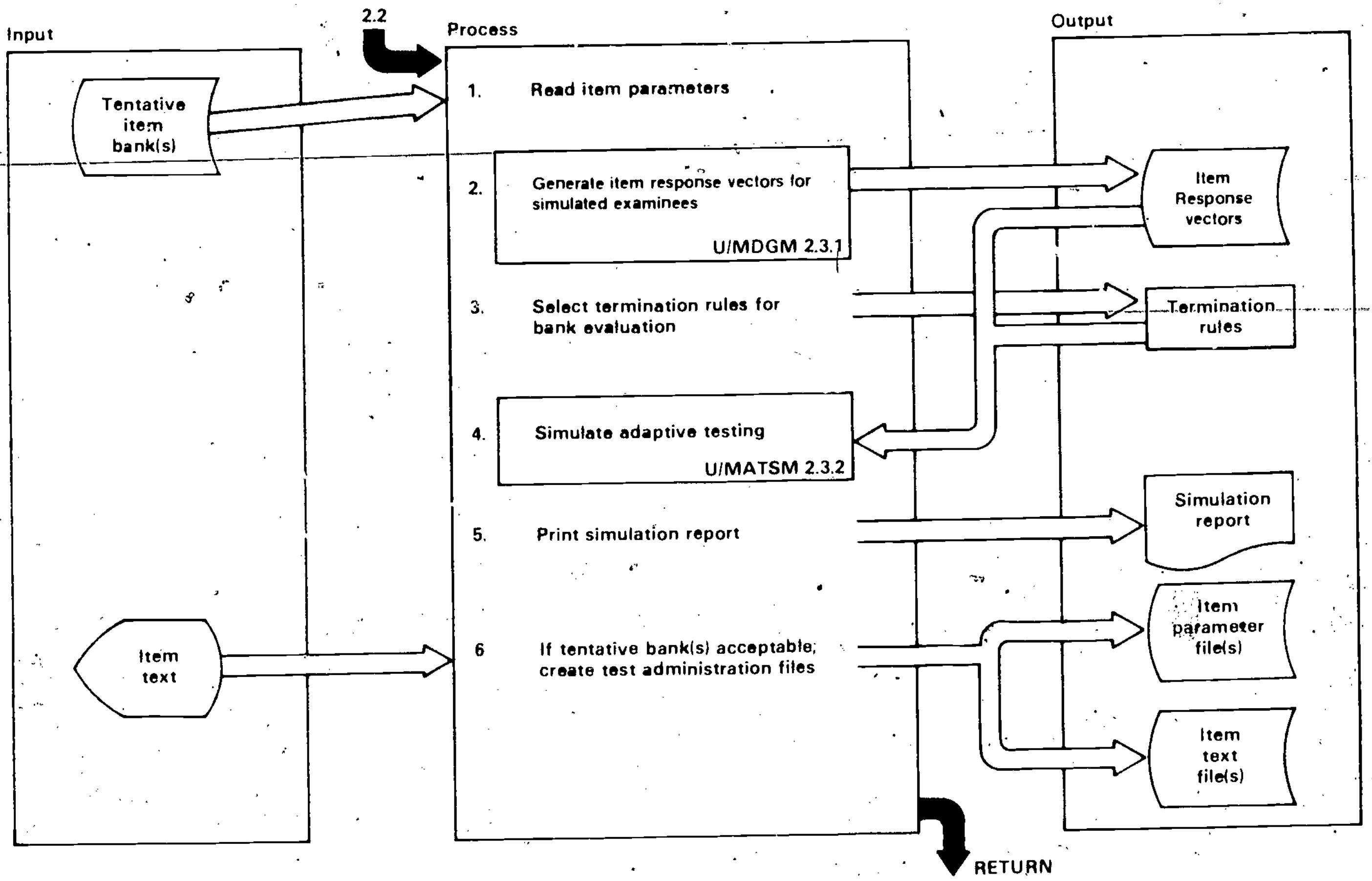

Figure 6. The item bank evaluation subfunction.

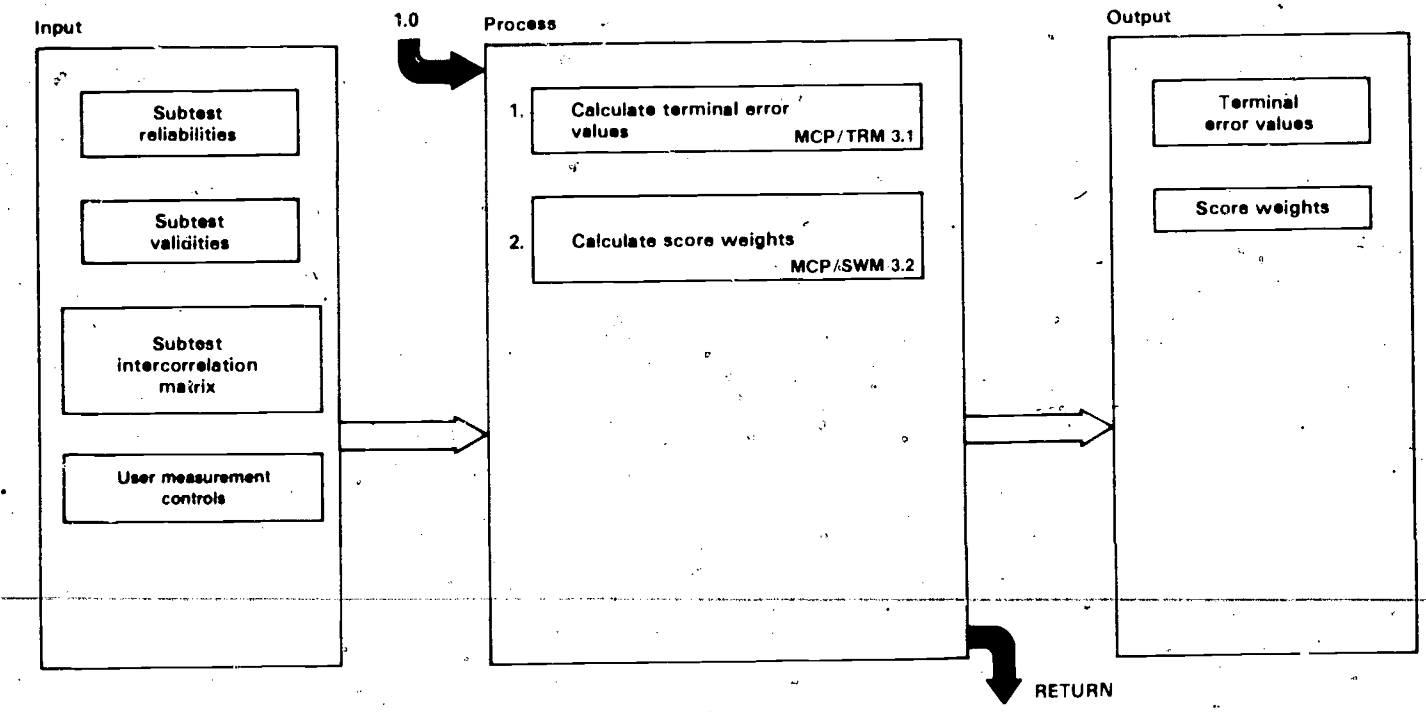

Figure 7. The measurement control function.

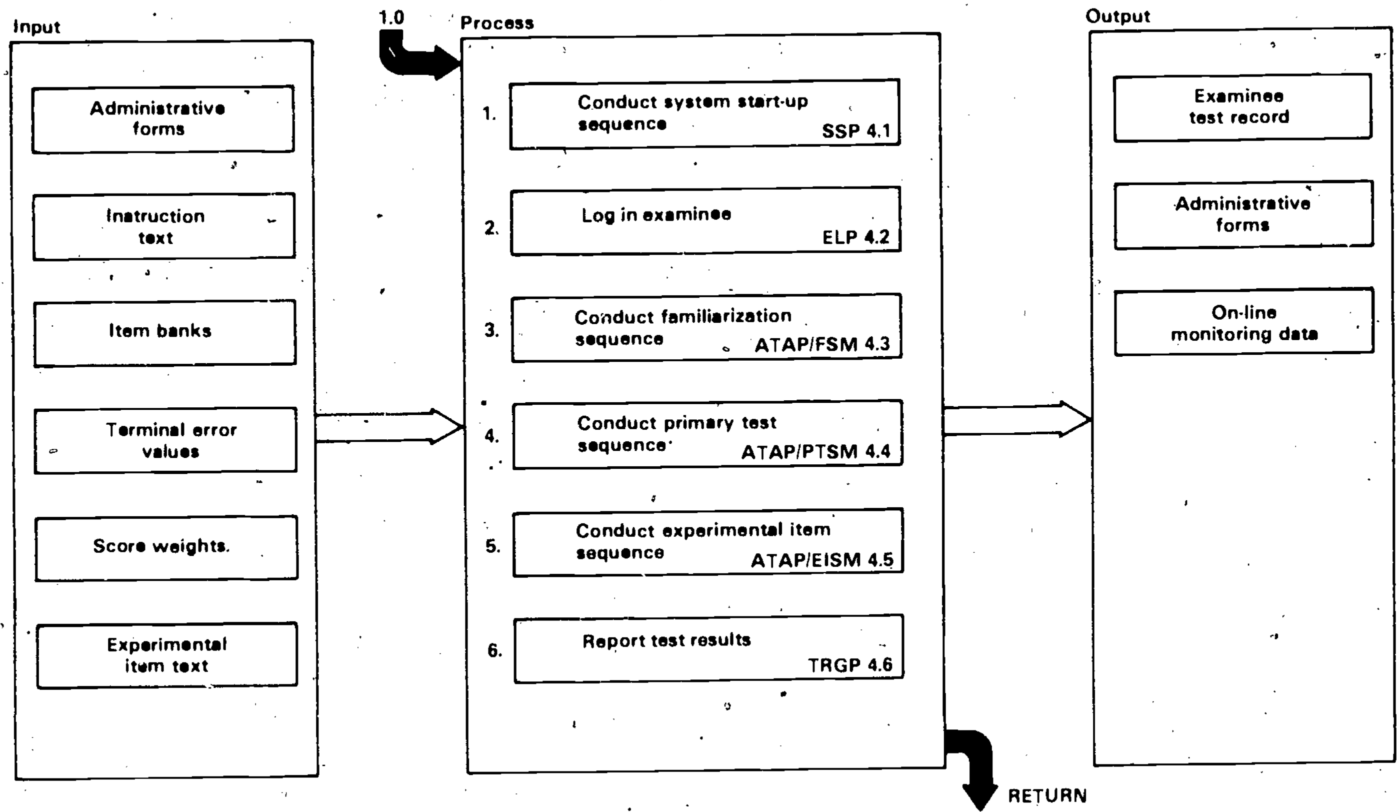

Figure 8. The test administration and scoring function.

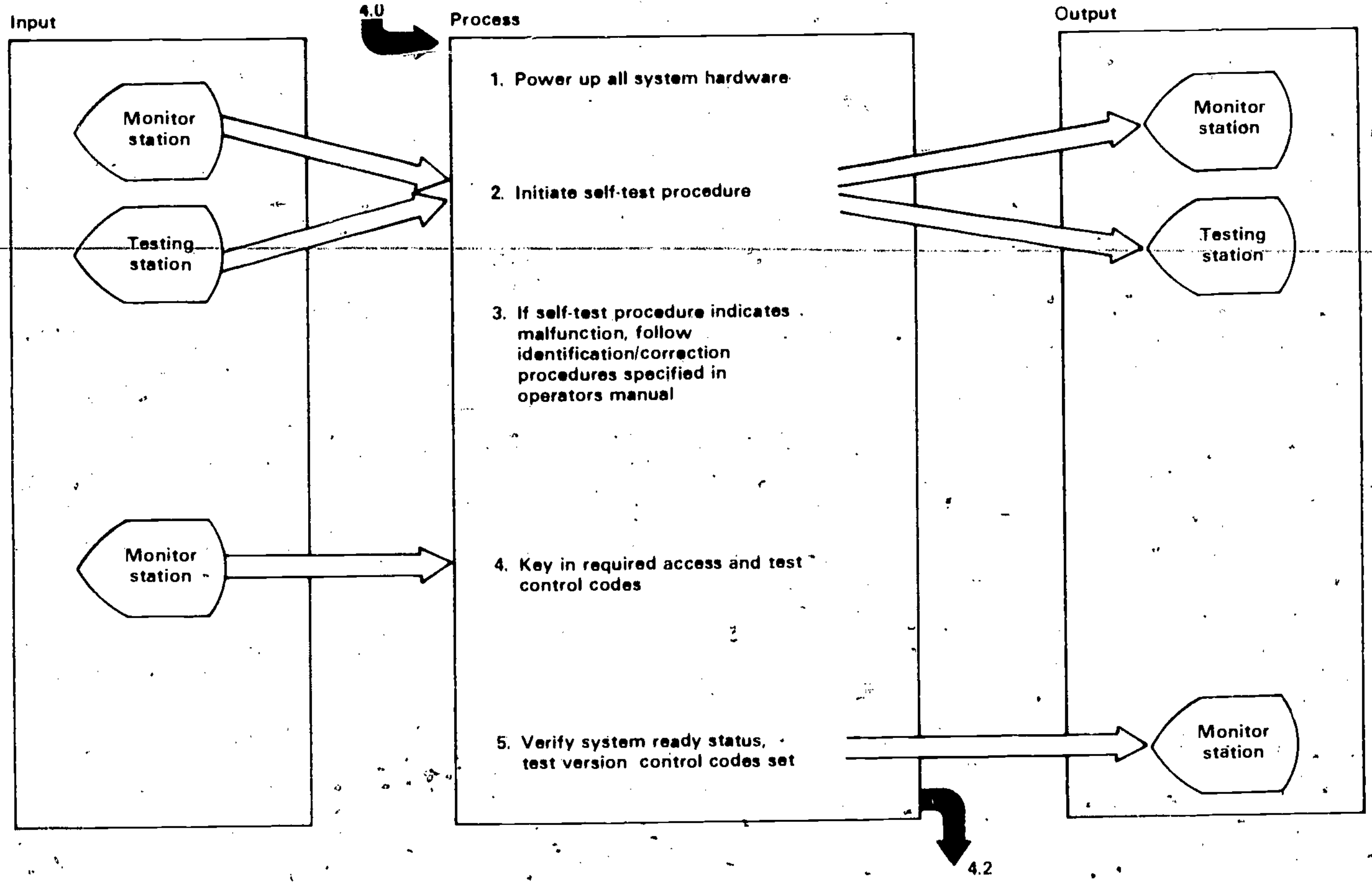

Figure 9. The system start-up subfunction.

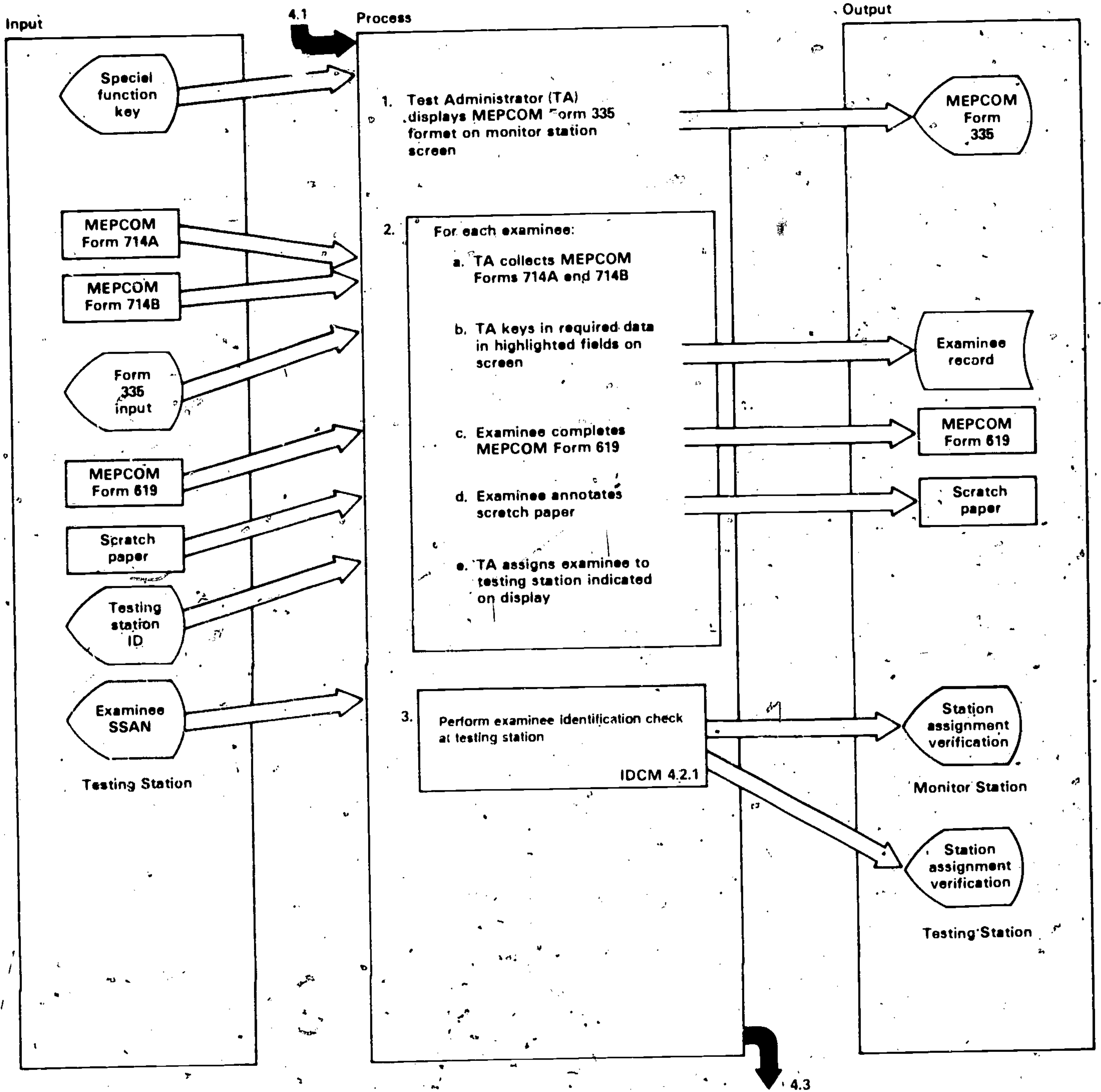

Figure 10. The examinee log-in subfunction.

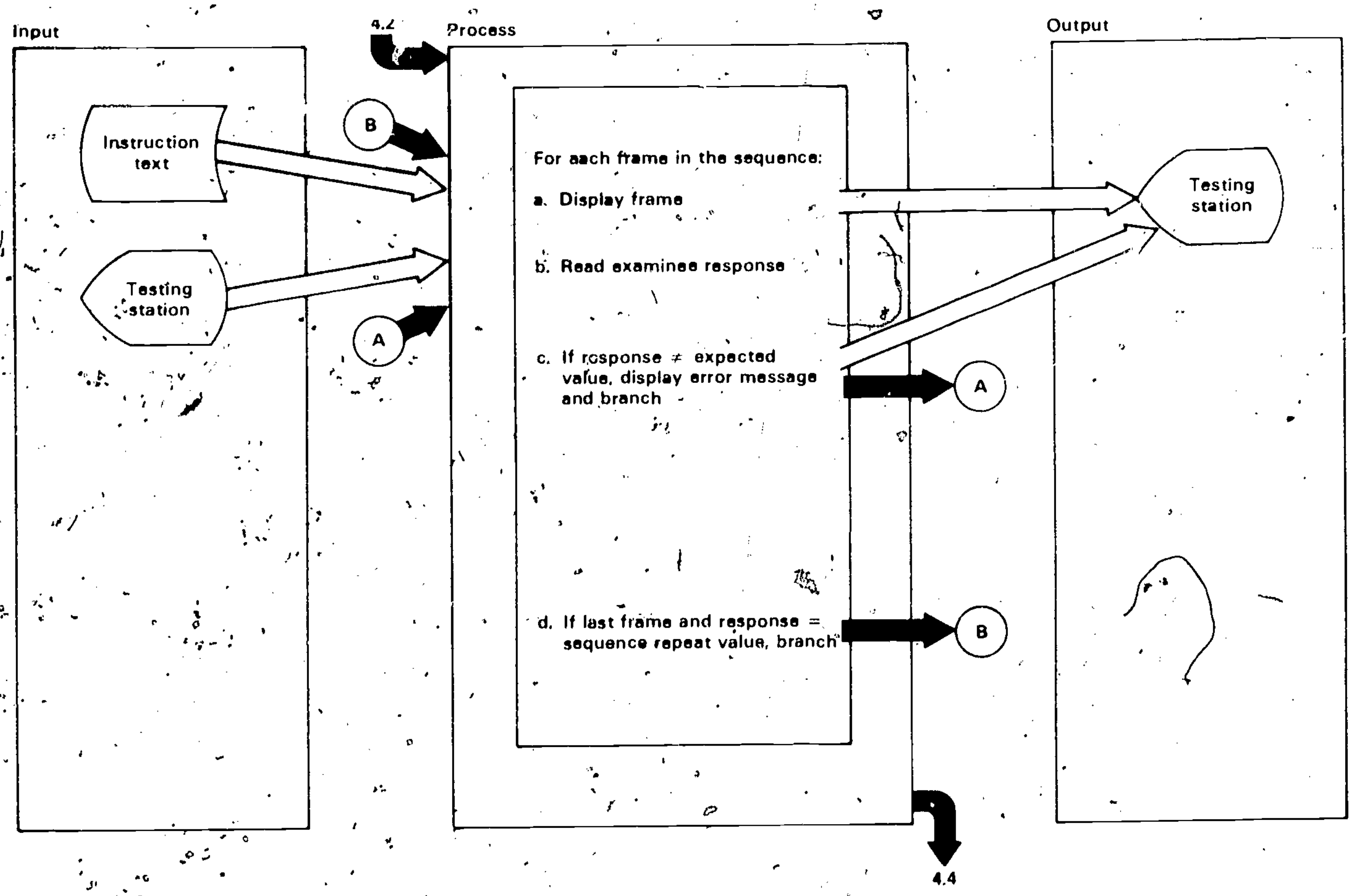

Figure 11. The familiarization subfunction.

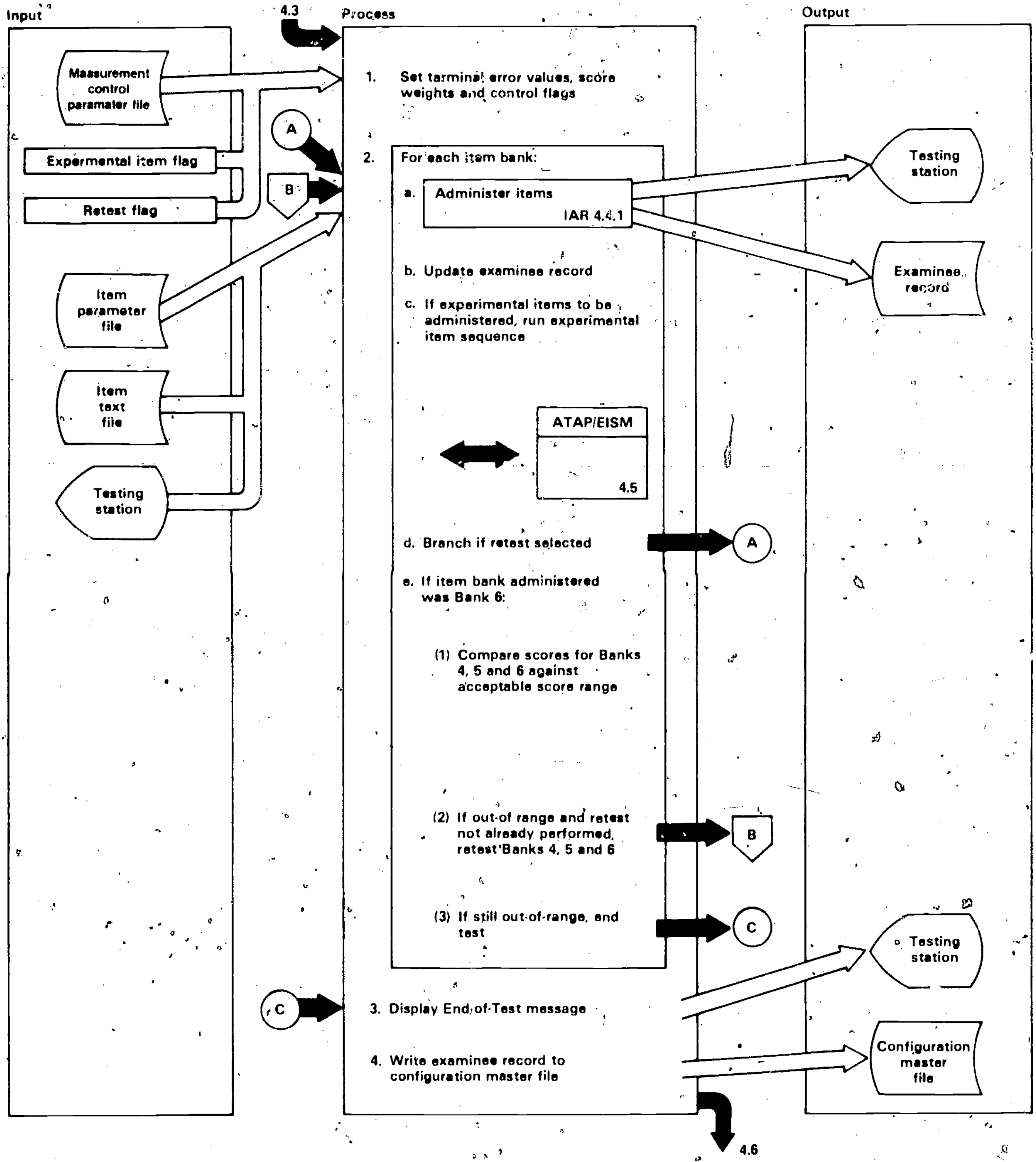

Figure 12. The primary test subfunction.

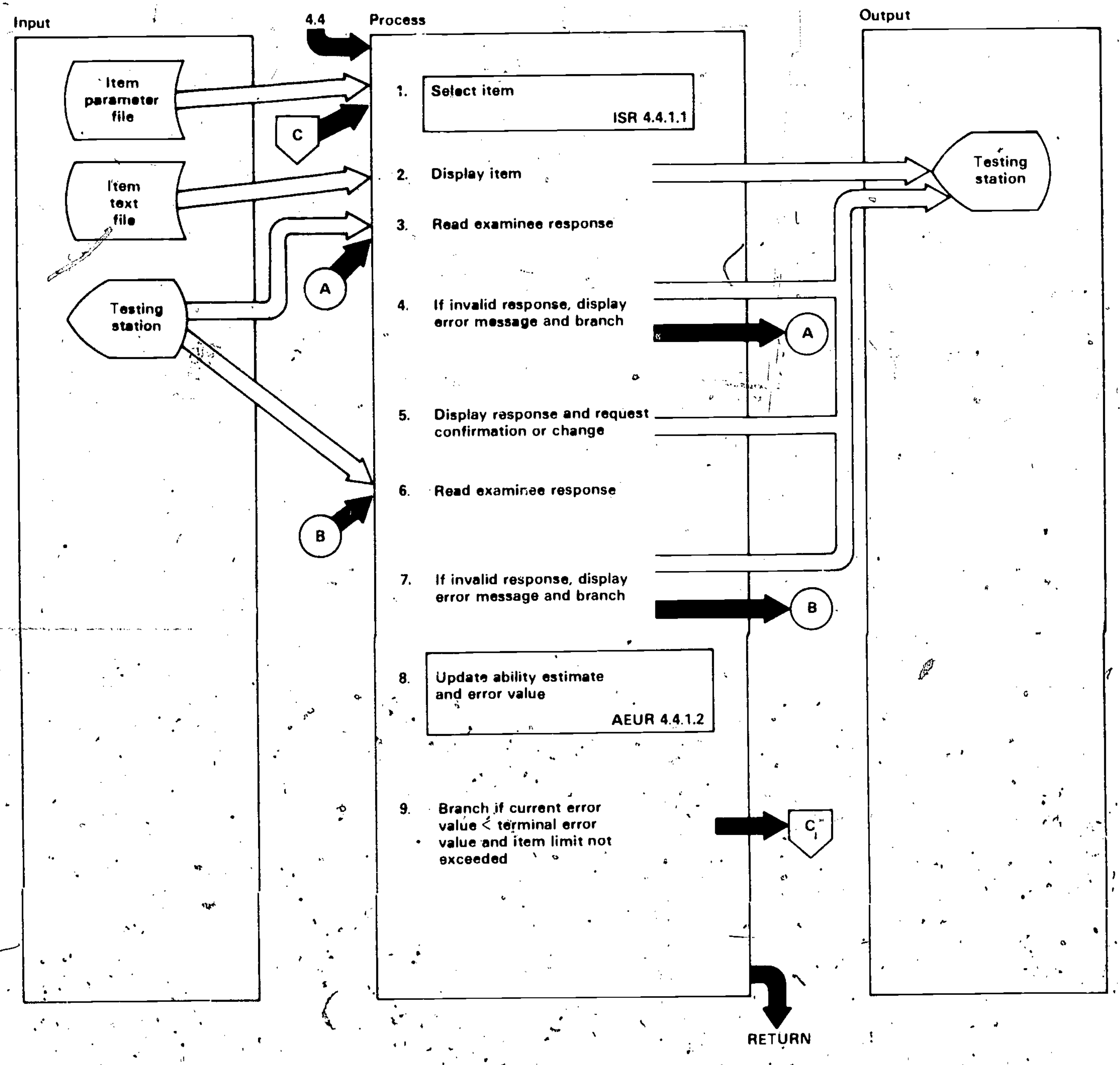

Figure 13. The test item administration subfunction.

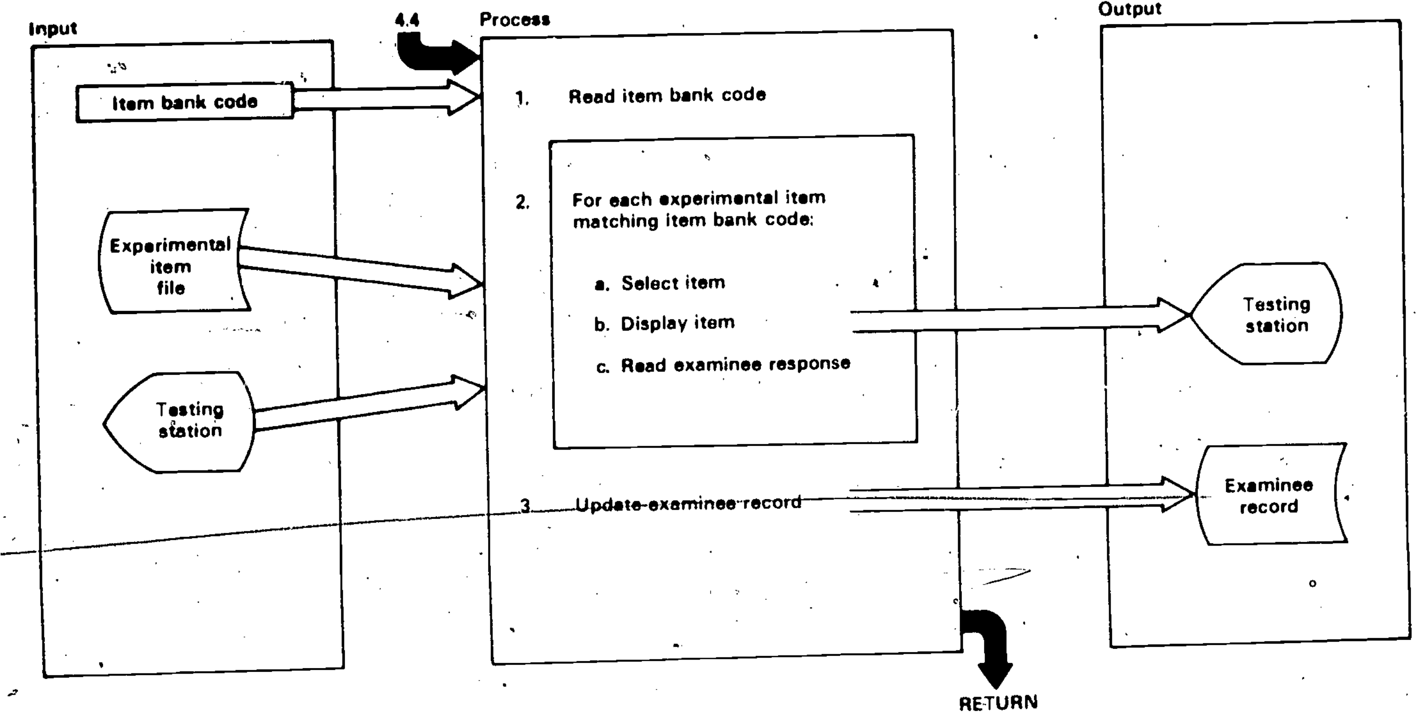

Figure 14. The experimental item subfunction.

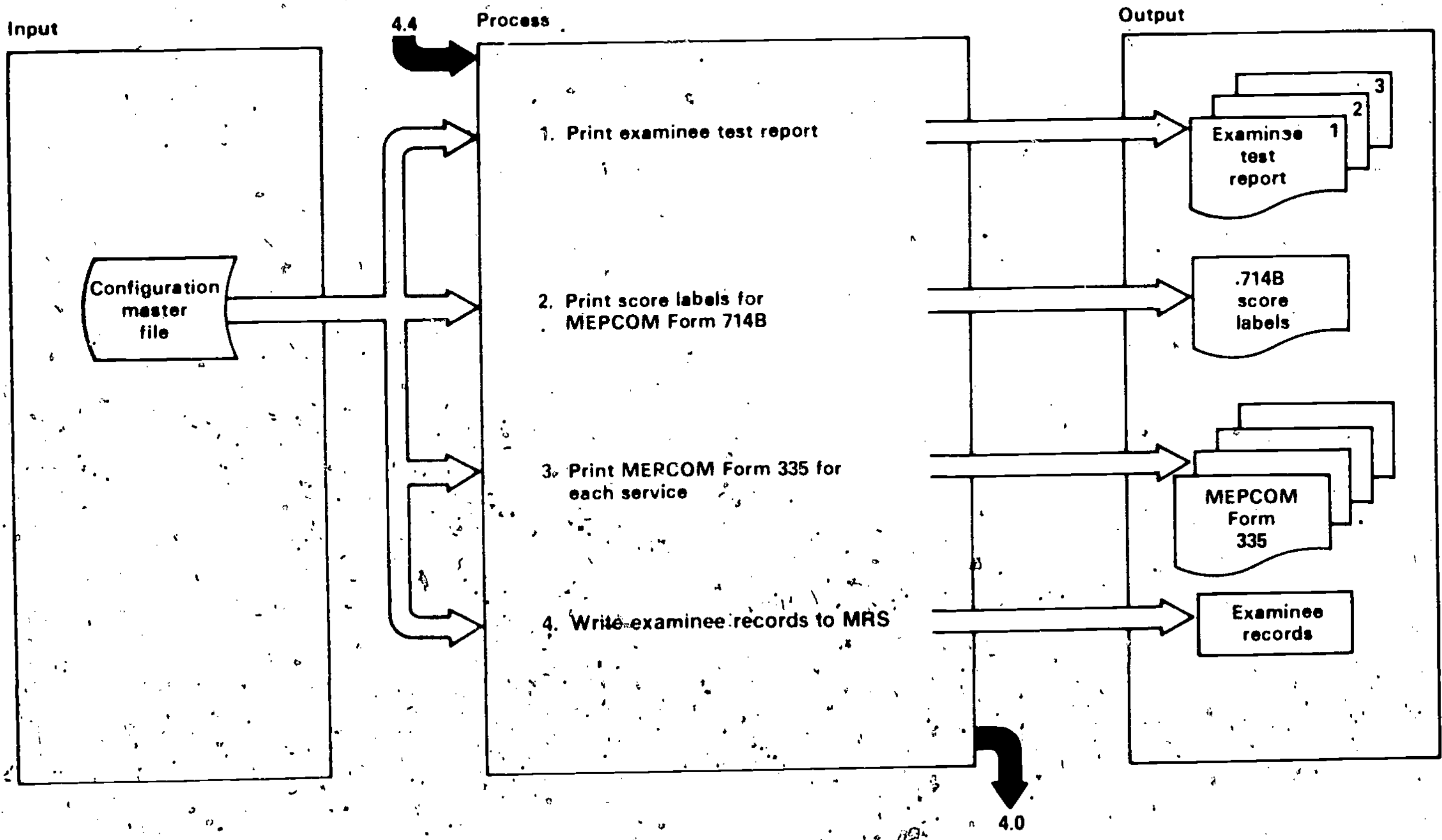

Figure 15. The test result reporting subfunction.

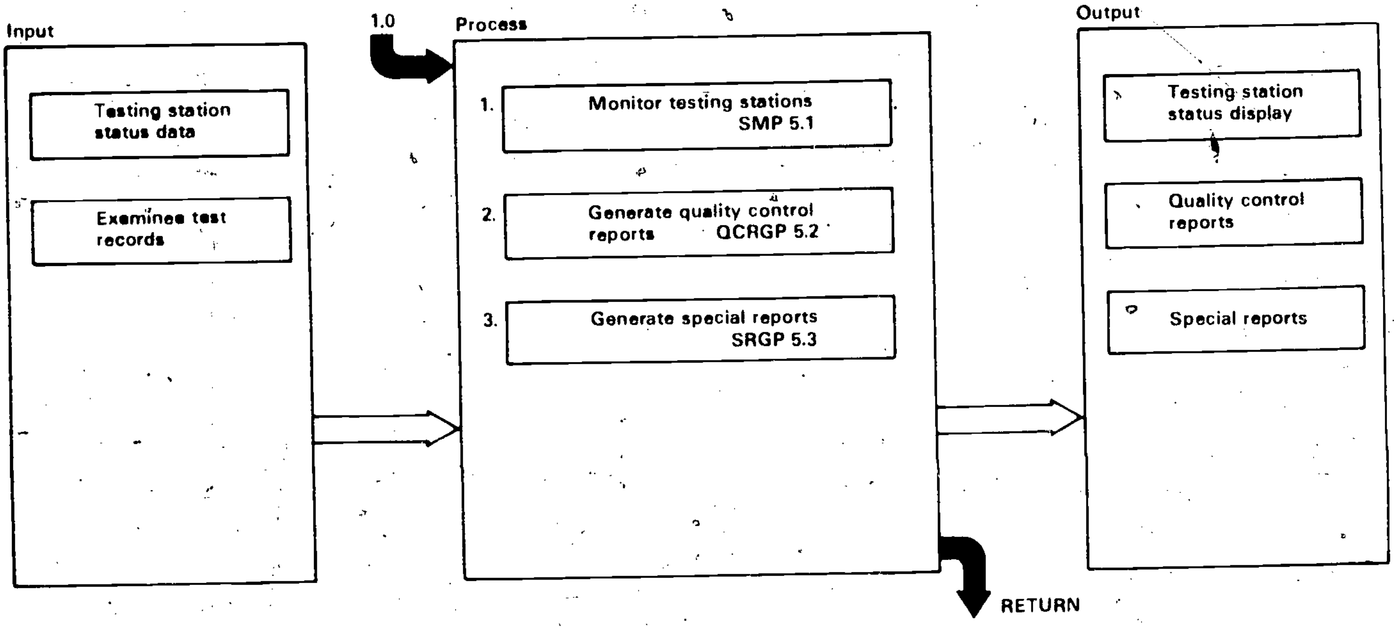

Figure 16. The monitoring and quality control function.

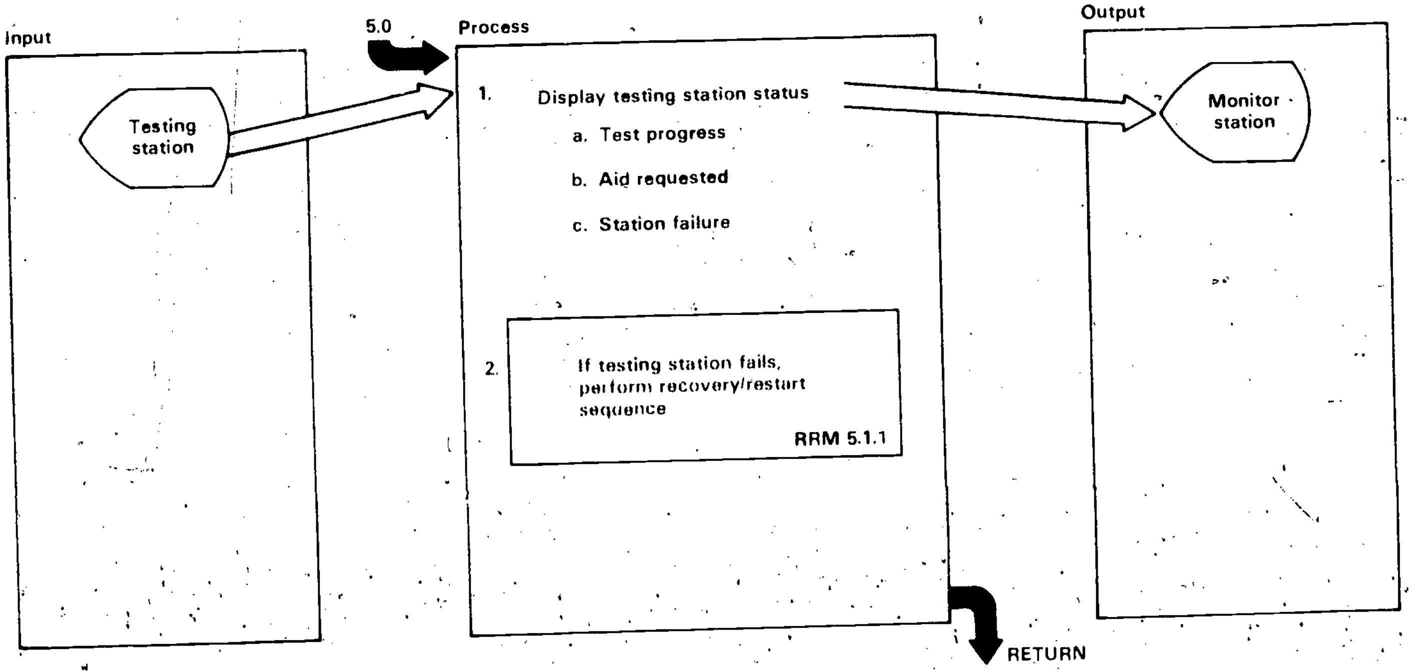

Figure 17. The testing station monitoring subfunction.

Item Banking Function

The CAT system's item banking function provides the sets of test questions, or item banks, necessary for adaptive test administration (Figure 3). It is composed of three subfunctions:

1. Test item calibration (Figure 4) refers to the estimation of the latent trait parameters, a_i, b_i, and c_i, of candidate test questions for item banking (Urry, 1981a). Input for this subfunction consists of results from either conventional or adaptive administration of the potential test questions. If parameters are to be estimated from conventional test results, examinee response data and scoring keys for the questions must be supplied. If parameters are to be estimated from adaptive test results, ability scores must be supplied as well. Algorithms for estimating parameters from conventional and adaptive test results have been described by Urry (1975, 1976, 1980) and Schmidt and Urry (1976). These algorithms are suggested as a guide for design of the CAT system's parameter estimation subfunctions. Parameter estimation from adaptive test results is especially important in CAT because it permits on-line calibration of potential test questions during normal operations. It provides a method for eventually ending dependence on conventional test results for item parameter estimation. The test item calibration subfunction produces parameter estimates and calibration statistics for the potential test questions. The parameter estimates are then treated as input to the item bank construction subfunction.

2. The item bank construction subfunction (Figure 5) takes the parameter estimates for candidate questions and compares them against target values for the a_i and c_i parameters. The prescription for acceptable values of these parameters has been detailed by Urry (1971, 1977b, 1981b). Questions that fail to meet this prescription are rejected by parameter values. The remaining item parameter sets are then sorted to ease later processing and a rectangular distribution of the items, by parameter, is built. Urry's prescriptions for the size and distributional shape of an item bank may be followed in selecting questions.

3. The item bank evaluation subfunction (Figure 6) is designed to assess the performance characteristics of an item bank before it is placed into operational use. It is one of the most critical quality control steps in CAT system design, because item bank performance characteristics are a major determinant of CAT system performance. A procedure for evaluating an item bank has been described by Urry (1974). From the functional perspective, the item parameter sets for the tentative item bank are used to generate response vectors (ones and zeros, or rights and wrongs) for simulated examinees. Termination rules are selected for item bank evaluation, based on the desired reliability of the bank (Urry, 1977b, 1981a). These rules are provided by specifying a value of the error of the ability estimate, at which point the test sequence is terminated. Adaptive testing is then simulated using the item parameter sets, response vectors, and termination rules. The results are reported. The item bank is made available, with associated question text, for operational use only if it is judged acceptable. The procedural steps in the item banking function are repeated for each ability for which an item bank is to be constructed. When several item banks will be administered as a multiple-ability battery, simulation of adaptive testing with the complete set of banks is conducted.

Measurement Control Function

The measurement control function, one of the most critical components of the CAT system, provides the means through which answers to the three basic questions underlying CAT are translated into system control parameters. These three questions are:

1. What is to be measured?

2. What degree of accuracy is to be employed?

3. How are subtest scores to be combined into composite scores?

User requirements are communicated to system personnel who, in turn, specify measurement protocols to meet the user's needs. These protocols embody the measurement requirements of each system user and determine both the way in which the adaptive testing process proceeds and the nature of its outputs. Furthermore, the protocols specify the combination of subtests required to meet specific measurement objectives (e.g., full ASVAB vs. Armed Forces Qualifications Test (AFQT) or service-specific composites), the outputs desired (e.g., subtest scores vs. weighted composite scores), and the scale and accuracy of measurement desired. They take the form of the input stream required by the system to generate control parameters.

It is through software generation of control parameters that user measurement protocols are implemented in the CAT system. These parameters are of three types: (1) termination rules, or terminal error values (values for the error of the estimate of ability), which determine the point in the adaptive testing sequence where testing for a particular ability is terminated, (2) subtest weights, which determine the relative contribution of a subtest score to a composite score (and which may be zero, if a subtest score is not to be included in a particular composite score), and (3) rescaling factors, which provide conversion of output scores based in the system's standard scale of measurement to scores based in an alternate scale of measurement.

The measurement control function must provide the capability for translation of a wide range of user measurement protocols into appropriate control parameters. The function can become complicated as the number and complexity of distinct user protocols increases. Its psychometric bases have been discussed by Urry (1980, 1981a & b). Its implementation depends on several necessary conditions of the total system design:

1. A Bayesian modal solution for item parameter estimates must be used.

2. The Owen-Bayesian algorithm must serve as the basis for item selection and ability estimation.

3. A variable-test-length termination strategy, based on target values of the standard error of the estimate of ability (for each subtest), must be employed.

A very simplified case of the measurement control function is illustrated in Figure 7.

Test Administration and Scoring Function

Administration and scoring of adaptive tests in the live testing environment (Figure 8) is often thought of as the sole function of a CAT system because it is the primary system function implemented in the field-resident physical system. It is composed of six subfunctions:

1. The system start-up subfunction (Figure 9) includes the steps necessary to prepare the physical system (the hardware and software) for a testing session. It includes power-up, self-test, sign-on, and system status verification activities.

2. The examinee log-in subfunction (Figure 10) performs the administrative tasks that identify the examinee to the system and that link the examinee's test record with the other steps in the applicant processing sequence. Inputs include data from administrative forms and examinee-supplied data, and outputs include administrative forms and the examinee record into which the test results will later be written. Additionally, a lower level subfunction has been specified to ensure that examinees are correctly seated at the testing stations to which they have been assigned.

3. The familiarization subfunction (Figure 11) is designed to familiarize the examinee both with the hardware and with the adaptive testing process. Introductory, instructional, and practice materials are displayed on the testing station display, and the examinee enters the required responses on the testing station keyboard. Checks are included to ensure that the examinee is proceeding through the familiarization sequence successfully. An option has also been designed for the examinee to request a repeat of the familiarization sequence. Inputs include introductory, instructional, and practice text, as well as examinee responses; outputs are displays of the input text and error messages.

4. The primary test subfunction (Figure 12) is the heart of the test administration and scoring function. It is designed to select and display test questions, read and score examinee responses, and update the examinee test record. It also provides administration of experimental items (through branching to another subfunction), selective retests, and test results recording on the testing site's master file. Inputs include control data, item parameters, item test, and examinee responses. Outputs include test item displays, error message displays, and the examinee test record.

Within the primary test subfunction, lower level subfunctions have been specified. The item administration subfunction (Figure 13) selects and displays test questions, reads examinee responses, and displays an error message when appropriate. It scores examinee responses and updates the estimate of ability and its associated error value. It terminates the testing sequence in a particular ability by checking the current error value of the ability estimate against the specified terminal error value. Because the item selection procedure and the ability and error updating procedures are psychometrically complex, lower level subfunctions for them have been identified but have not been specified in separate HIPO diagrams. Decisions about these subfunctions will have to be made within the context of the system's psychometric development activities. Urry (1977b, 1980, 1981a & b) has offered guidance in developing these procefures.

5. The experimental item subfunction (Figure 14) provides administration of experimental, or potential, test questions within the context of an adaptive test. It selects and displays experimental items, and reads and records examinee responses. Inputs include item bank codes, item text, and examinee responses; outputs include item text displays and examinee responses to the items. This subfunction is called by primary test subfunction when control codes indicate that experimental items are to be administered.

6. The test results reporting subfunction (Figure 15) is designed to provide printed reports of test results, including any required administrative forms. It inputs data from the testing site's configuration master file and prints reports as required. It is also designed to feed testing results into the MEPS reporting system.

Monitoring and Quality Control Function

This component, which provides system-wide quality control of all CAT system functions as well as monitoring of the on-site testing process, is composed of three subfunctions: testing station monitoring, quality control report generation, and special report generation (Figure 16). The term "quality control," as used in this function, implies not only physical system diagnostics and maintenance but also monitoring and control of the psychometric integrity of the CAT system. Because the system will stand or fall on the quality of its personnel measurement, its psychometric integrity requires constant scrutiny.

The testing station monitoring subfunction (Figure 17) may be used in various ways. During a testing session, three conditions might occur that would require the attention of the test monitor: (1) The examinee might fail to progress normally through the testing sequence and also fail to request assistance, (2) the examinee might, for any reason, request monitor assistance, or (3) a failure might occur in a testing station. The testing station monitoring subfunction should provide a constant display of testing station status, so that such conditions may be identified. Additionally, if a testing station fails, a lower level subfunction should be initiated to perform a recovery and restart sequence. Because this lower level subfunction is dependent on decisions yet to be made about the nature of the recovery and restart procedures desired for the CAT system, it has not yet been specified in this HIPO package.

CAT System Structure

The task of the system designer is to define system functions and to translate those functions into structure, logic, and organization--the set of design specifications used in the system development stage. Bingham and Davies (1972) list 15 main activities in the development of a detailed system design for implementation. These activities include development of comprehensive design documentation, as well as final specification of all inputs and outputs, data and control paths, file structures, overall system logic, software and hardware, and internal and external interfaces. CAT system structure consists of the concrete elements (Ackoff, 1974) required to implement system functions in the real world. The Bingham and Davies activities suggest the type of concrete elements with which the system designer must be concerned.

The four major functions identified in the CAT functional design model suggest a system structure that implements each function in a separate subsystem with its own data, logic, hardware, and software characteristics. Modular design concepts, applied to separating system functions into concrete subsystems and to developing the concrete elements of those subsystems, allow the system to evolve gracefully in step with changes in operational requirements or the availability of new technology. The following discussion of CAT system structure is an example of translation of the functional design model into such concrete system elements. The discussion focuses on system software specification because the functional design is primarily embodied in such software. Table 1 presents system software components by system function.

Item Banking Subsystem

The item banking function described in the functional design model is implemented by the item banking subsystem (IBS), a structural component that consists of three major computer programs. These programs contain eight software modules with associated file structures, control logic, and interfaces. They interface with each other through their file structures and with the rest of the system by providing item bank files to the test administration and scoring subsystem.

Table 1

CAT System Software Components, Enumerated by System Function

System Function	Software Component			
	Subsystem	Program	Module	Subroutine
1.0 CAT System Overview	--	--	--	--
2.0 Construct, test, and evaluate item banks	Item banking (IBS)	--	--	--
2.1 Calibrate test items	--	Test calibration (TCP)	--	--
2.1.1 Calculate parameter estimates from conventional test results	--	--	Conventional test calibration (CTCM)	--
2.1.2 Calculate parameter estimates from adaptive test results	--	--	Adaptive test calibration (ATCM)	--
2.2 Construct item banks	--	Item bank construction (IBCP)	Item sort (ISM)	--
2.2.1 Build rectangular item distribution	--	--	Rectangular item distribution (RIDM)	--
2.3 Evaluate bank performance	--	Item bank evaluation (IBEP)	--	--
2.3.1 Generate item response vectors	--	--	Univariate data generator (UDGM) Multivariate data generator (MDGM)	--
2.3.2 Simulate adaptive testing	--	--	Univariate adaptive testing simulation (UATSM) Multivariate adaptive testing simulation (MATSM)	--
3.0 Generate measurement control parameters	Measurement control (MCS)	Measurement control (MCP)	--	--
3.1 Calculate terminal error values	--	--	Termination rule (TRM)	--
3.2 Calculate score weights	--	--	Score weighting (SWM)	--

Table 1 (Continued)

System Function	Software Component			
	Subsystem	Program	Module	Subroutine
4.0 Administer and score adaptive tests	Test administration and scoring (TASS)	--	--	--
4.1 Perform system start-up procedure	--	System start-up (SSP)	Self-test (STM)	--
4.2 Log in examinee	--	Examinee log-in (ELP)	--	--
4.2.1 Perform examinee identification check	--	--	Identification check (IDCM)	--
4.3 Conduct familiarization sequence	--	Adaptive test administration (ATAP)	Familiarization sequence (FSM)	--
4.4 Conduct primary test sequence	--	--	Primary test sequence (PTSM)	--
4.4.1 Administer items	--	--	--	Item administration (IAR)
4.4.1.1 Select item	--	--	--	Item selection (ISR)
4.4.1.2 Update ability estimate and error value	--	--	--	Ability error update (AEUR)
4.5 Conduct experimental item sequence	--	--	Experimental item sequence (EISM)	--
4.6 Report test results	--	Test report generator (TRGP)	--	--
5.0 Monitor system performance; provide quality control reports	Monitoring/quality control (MQCS)	--	--	--
5.1 Monitor testing stations	--	Station monitoring (SMP)	--	--
5.1.1 Perform recovery/restart procedure	--	--	Recovery/restart (RRM)	--
5.2 Generate quality control reports	--	Quality control report generator (QCRGP)	--	--
5.3 Generate special reports	--	Special report generator (SRGP)	--	--

1. The test calibration program (TCP) calibrates potential test questions, using input from either conventional or adaptive test results, and writes calibration results to a parameter estimate file. It also prints a report of the calibration process. Two software modules actually perform the item parameter estimation functions: The conventional test calibration module (CTCM) calculates parameter estimates and calibration statistics from conventional test results, and the adaptive test calibration module (ATCM) performs the calculations from adaptive test results. Required files include (a) a control card file consisting of program control parameters, item labels, and item keys, (b) a file containing conventional test results, including item response data, (c) a file containing adaptive test results, including examinee item response data and ability scores, and (d) a file into which item parameter estimates will be written.

2. The item bank construction program (IBCP) reads the parameter estimate file, rejects item parameter sets that do not meet the prescription for values of the a_i and c_i parameters, sorts the remaining sets, and builds a rectangular distribution of those sets by b_i values. Those item parameter sets are written to a file as the tentative item bank, and a bank composition report is printed. The item sort module (ISM) performs the item sorting task, and the rectangular item distribution module (RIDM) performs the task of building the rectangular item distribution from the sort results. Required files include a parameter estimate file, a file into which the item sort results are written, a file containing the rectangular item distribution, and a file to contain the tentative item bank.

3. The item bank evaluation program (IBEP) reads the parameter sets contained in the tentative item bank, generates response vectors for simulated examinees, and applies the termination rules selected for bank evaluation to simulate adaptive testing with the tentative item bank. It prints a report of the simulation process and creates the item bank files required for test administration. When multiple banks are to be used as a test battery, response vectors are generated and adaptive testing is simulated for the set of item banks as well. The univariate data generator module (UDGM) generates response vectors for single bank evaluation, and the multivariate data generation module (MDGM) performs the same task for multiple bank evaluation. The univariate adaptive testing simulation module (UATSM) simulates adaptive testing with a single item bank, while the multivariate adaptive testing simulation module (MATSM) simulates it with multiple item banks. Required files include a tentative item bank or banks, a file containing generated response vectors, a file (or files) to contain text for the items in the operational bank, and a file (or files) to contain the parameters for those items. Termination rules and item text must be supplied as additional data.

Measurement Control Subsystem

Because the measurement control function cannot be adequately specified until the range of user requirements has been defined, some structural elements can only be suggested. The measurement control subsystem (MCS) will consist of several software components, of which the measurement control program (MCP), containing two modules, is only illustrative. The termination rule module (TRM) calculates termination rules for either single- or multiple-ability adaptive tests, and the score weighting module (SWM) calculates score weights to be applied in developing a multiple-ability composite score. Files required are a file containing subtest reliabilities and validates, a file representing the subtest intercorrelation matrix, and a file into which terminal error values and score weights will be written. Data representing user measurement protocols are also required as input to the program. This subsystem interfaces with the remainder of the system by providing measurement control parameters (terminal error values and score weights) to the test administration and scoring subsystem.

Test Administration and Scoring Subsystem

The test administration and scoring subsystem (TASS) comprises the major portion of the CAT system functional design model. It consists of four computer programs, five modules, and three subroutines, plus associated file structures, data requirements, control logic, and interfaces.

1. The system start-up program (SSP), upon system power-up, readies the hardware configuration at the testing site for the start of a testing session. The SSP includes a self-test module (STM) that performs an automatic check of system hardware and signals when the system is ready for operation. The program reads access and test control codes from the test monitor station and verifies system status on the station's display. When system-ready status is indicated, the SSP passes control to the examinee log-in program.

2. The examinee log-in program (ELP) displays a data entry format for the test monitor, reads identification data entered by the test monitor for each examinee, and creates the examinee record. The identification check module (IDCM) verifies that examinees are seated at the testing stations to which they have been assigned. This program requires a file into which the examinee records will be written. When examinee placement at a testing station has been verified, the program passes control to the adaptive test administration program.

3. The adaptive test administration program (ATAP) implements the familiarization, primary test, and experimental item subfunctions of the model. The familiarization sequence is conducted by the familiarization sequence module (FSM), which displays each frame in the sequence on the testing station display, reads examinee responses, and checks to see whether the responses match expected values. It will also initiate a repeat of the sequence if the response to the last frame matches a specified value. Upon completion of the familiarization sequence, the module passes control to the primary test sequence module (PTSM). After reading termination and weighting control data and experimental item and selective retest flags, the PTSM conducts the primary test sequence for each item bank to be administered. It administers items, updates the examinee record, branches to the experimental item sequence module if experimental items are to be administered, conducts a retest with an item bank when required, and terminates the test, writing the examinee record into the testing site's configuration master file. When required, it conducts a retest with the AFQT portion of the ASVAB and then proceeds with testing or terminates the test at the point, depending on the outcome of the retest.

Several functions of the PTSM are implemented in subroutines. The item administration subroutine (IAR) displays test questions, reads examinee responses, checks response validity, and displays error messages. The IAR also checks the current error value of the estimate of examinee ability against the specified terminal error value. It checks to see whether a specified limit for the number of items to be administered in any one bank has been exceeded. This subroutine passes control to the item selection subroutine (ISR) for test question selection and to the ability and error update subroutine (AEUR) for the scoring of examinee responses and updating of ability and error estimates.

For administration of experimental items, control is passed to the experimental item sequence module (EISM), which reads the current item bank code and selects and displays experimental test questions. It also reads examinee responses to the questions and records those responses in the examinee record. It then passes control back to the PTSM.

4. The test report generator program (TRGP) reads the test site's configuration master file and prints examinee test reports and administrative forms when they are required. It also writes examinee records into the MEPS reporting system through that system's interface with the monitor station. Program control is initiated by the test monitor through the monitor station keyboard.

File requirements for the subsystem include (1) a file into which the examinee records will be written, (2) a file containing introductory, instructional, and practice text, (3) the termination and weighting control file, (4) the item bank parameter and text files, (5) an experimental item file, and (6) the configuration master file. Data requirements include system access and control codes, examinee identification data, experimental item and selective retest control flags, and examinee responses. The programs in this subsystem interface with each other through their internal control structures and through the subsystem's file structure. The subsystem interfaces with the remainder of the CAT system through the overall system file structure and through direct data and control links with the monitoring and quality control subsystem.

Monitoring and Quality Control Subsystem

Three programs constitute the monitoring and quality control subsystem. At the test monitor station, the station monitoring program (SMP) provides a display of testing status, including test progress, aid requested, station failure, and system problems (e.g., psychometric anomalies). It also includes a recovery and restart module (RRM) to initiate a recovery and restart sequence in the event of testing station failure. The quality control report generator program (QCRGP) analyzes systemwide performance data and prints quality control reports, as required. The special report generator program (SRGP) provides special analyses of system performance data and subsequently generates reports based on those analyses. File requirements for this subsystem would include access to all CAT system permanent files and the generation of any analysis files required. Data requirements primarily include testing station status data. Interfaces to the remainder of the CAT system are accomplished through the system's file structure, except for the station monitoring program, which requires direct data and control links to the test administration and scoring subsystem.

CAT System Implementation

Hardware

System hardware must support two categories of system functions: (1) those implemented within the context of the actual testing situation (i.e., at the test site), and (2) those implemented elsewhere (i.e., at a laboratory or administrative headquarters). A testing site may be a permanent location, such as a MEPS, or a temporary location, such as a high school or a local post office. Thus, the choice of hardware and the determination of the way in which that hardware is configured present a complicated problem. Table 2 displays system functions in comparison to hardware functions. System mode, processing, input/output, and storage requirements have been indicated for each function and subfunction in the CAT system functional design model. Categories of hardware that might satisfy those requirements have also been indicated. These categories are generic and include medium-to-large-scale mainframe systems, small-to-medium-scale minicomputers, microprocessors, hard disks, floppy disks, alphanumeric displays, graphics displays, keyboards, and printers. Making these hardware choices will require careful consideration on the part of system designers; the task goes beyond the realm of the preliminary design considerations discussed here. However, the issue of hardware support at the testing site deserves preliminary consideration in light of recent advances in microcomputer technology.

Table 2

CAT Hardware Functions, Enumerated by System Function

System Function	Hardware Function																													
	System Mode			Processing												Category	I/O								Category	Storage				Category
				Arithmetic				Input/ Output			Communications			File Ops			Input			Output										
	Batch	Real Time	Real Time Time Critical	Complex Equations	Floating Point	Iteration	Matrix Manipulation	File	Keyboard Display	Printer	Bus-Level	Local Serial	Remote Serial	Multiple Sort	On-line Edit		SASD	DASD	Keyboard	SASD	DASD	Alphanumeric Display	Graphics Display	Hard Copy		Less than 500 KB	500 KB – 1 MB	1 MB – 5 MB	5 MB – 10 MB	
1.0 CAT System Overview																														
2.0 Construct, test and evaluate item banks	X	–	–	X	X	X	X	X	X	X	–	–	–	X	X	A	X	X	X	X	X	X	–	X	A	–	–	–	X	A
2.1 Calibrate test items	X	–	–	X	X	X	–	X	–	X	–	–	–	–	–	–	X	X	–	X	X	–	–	X	–	X	–	–	–	–
2.1.1 Calculate parameter estimates from conventional test results	X	–	–	X	X	X	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	X	–	–	–	–
2.1.2 Calculate parameter estimates from adaptive test results	X	–	–	X	X	X	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	X	–	–	–	–
2.2 Construct item banks	X	–	–	X	X	X	–	X	–	X	–	–	–	X	–	–	X	X		X	X	–	–	X	–	X	–	–	–	–
2.2.1 Build rectangular item distribution	X	–	–	X	X	X	–	X	–	–	–	–	–	X	–	–	X	X		X	X	–	–	X	–	X	–	–	–	–
2.3 Evaluate bank performance	X	X	–	X	X	X	X	X	X	X	–	X	X	–	X	–	X	X	X	X	X	X	–	X	–	–	–	–	X	–
2.3.1 Generate item response vectors	X	–	–	X	X	X	X	X	–	X	–	–	–	–	–	–	X	X	–	X	X	–	–	X	–	X	–	–	–	–
2.3.2 Simulate adaptive testing	X	–	–	X	X	X	X	X	–	X	–	–	–	–	–	–	X	X	–	X	X	–	–	X	–	–	X	–	–	–
3.0 Generate measurement control parameters	X	–	–	X	X	X	X	X	–	X	–	–	–	–	–	A	X	X	–	X	X	–	–	X	A	X	–	–	–	A
3.1 Calculate terminal error values	X	–	–	X	X	X	X	X	–	X	–	–	–	–	–	–	X	X	–	X	X	–	–	X	–	X	–	–	–	–
3.2 Calculate score weights	X	–	–	X	X	X	X	X	–	X	–	–	–	–	–	–	X	X	–	X	X	–	–	X	–	X	–	–	–	–
4.0 Administer and score adaptive tests	–	X	X	X	X	X	–	X	X	X	X	X	–	–	–	B,C	–	X	X	–	X	X	X	X	D,E	–	–	X	–	D,E
	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	F,G	–	–	–	–	–
	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	H,I	–	–	–	–	–
4.1 Perform system start-up sequence	–	X	–	–	–	–	–	X	X	X	X	X	–	–	–	–	–	X	X	–	X	X	X	X	–	X	–	–	–	–
4.2 Log in examinee	–	X	–	–	–	–	–	X	X	–	X	X	–	–	–	–	–	X	X	–	X	X	–	–	–	–	–	X	–	–
4.2.1 Perform examinee identification check	–	–	X	–	–	–	–	–	X	–	X	X	–	–	–	–	–	–	X	–	–	X	–	–	–	–	–	–	–	–
4.3 Conduct familiarization sequence	–	–	X	–	–	–	–	X	X	–	X	–	–	–	–	–	–	X	X	–	–	X	–	–	–	X	–	–	–	–
4.4 Conduct primary test sequence	–	–	X	X	X	X	–	X	X	–	X	X	–	–	–	–	–	X	X	–	X	X	X	–	–	–	–	X	–	–
4.4.1 Administer items	–	–	X	X	X	X	–	X	X	–	X	–	–	–	–	–	–	X	X	–	X	X	X	–	–	–	–	X	–	–
4.4.1.1 Select item	–	–	X	X	X	X	–	X	X	–	–	–	–	–	–	–	–	X	–	–	X	–	–	–	–	–	–	–	–	–
4.4.1.2 Update ability estimate and error value	–	–	X	X	X	X	–	X	–	–	–	–	–	–	–	–	–	–	–	–	X	–	–	–	–	X	–	–	–	–
4.5 Conduct experimental item sequence	–	–	X	–	–	–	–	X	X	–	X	–	–	–	–	–	–	X	X	–	X	X	X	–	–	X	–	–	–	–
4.6 Report test results	–	X	–	–	–	–	–	X	X	X	–	–	X	–	–	–	–	X	X	–	X	X	–	X	–	–	X	–	–	–
5.0 Monitor system performance, provide quality control reports	X	X	–	–	–	–	–	X	X	X	X	X	X	–	–	B,C	X	X	X	X	X	X	–	X	D,I H,I	–	–	–	X	D
5.1 Monitor testing stations	–	X	–	–	–	–	–	X	X	–	X	X	–	–	–	–	–	X	X	–	–	X	–	–	–	–	–	–	X	–
5.1.1 Perform recovery/restart sequence	–	X	X	–	–	–	–	X	X	X	X	X	–	–	–	–	–	X	X	–	X	X	–	X	–	–	–	–	X	–
5.2 Generate quality control reports	X	–	–	–	–	–	–	X	–	X	–	–	X	–	–	–	X	X	X	X	X	–	–	X	–	–	–	–	X	–
5.3 Generate special reports	X	–	–	–	–	–	–	X	–	X	–	–	X	–	–	–	X	X	X	X	X	–	–	X	–	–	–	–	X	–

Hardware Category A Medium to large scale mainframe system, B Small to medium scale minicomputer, C Microprocessor, D Hard disk,
E Floppy disk F Alphanumeric display, G Graphics display, H Keyboard, I Printer

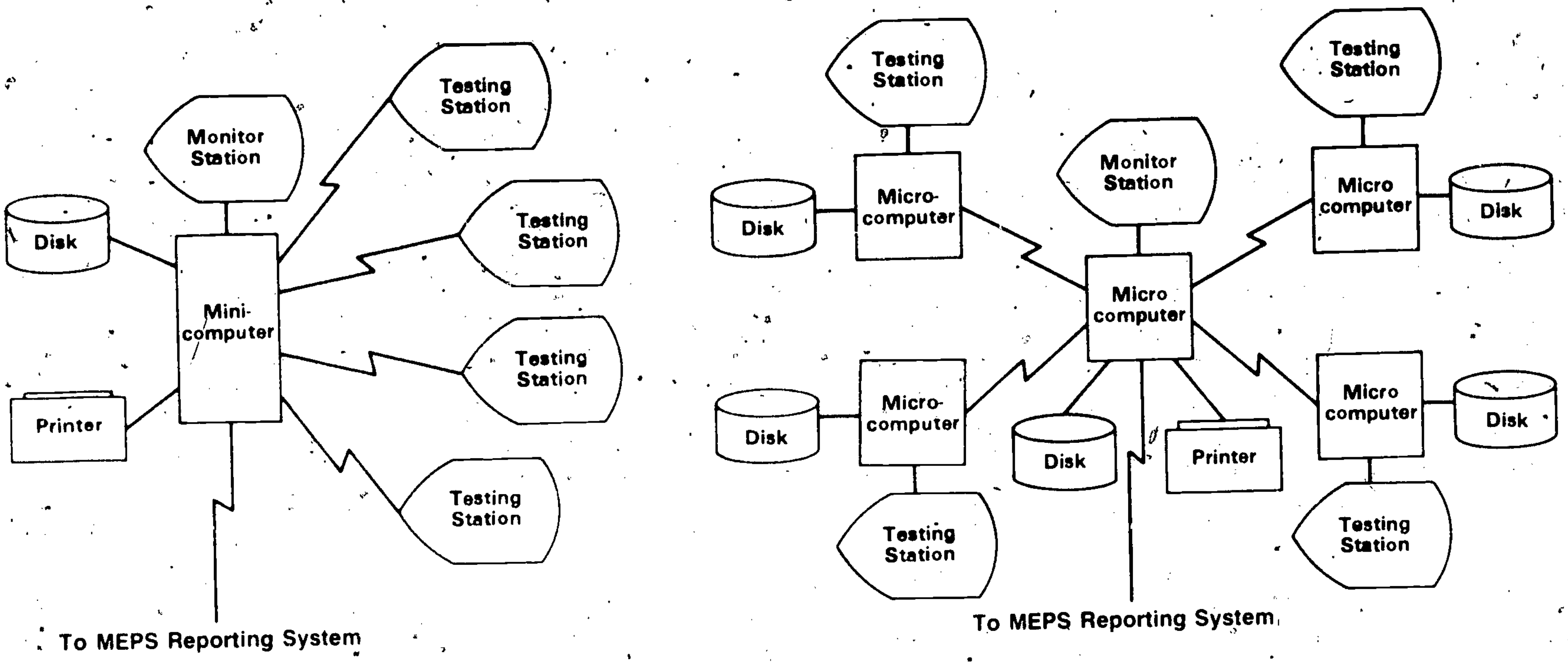

Figure 18a. **Minicomputer-Based Configuration**

Technology:	**Straightforward application of well-established technology**
Functional Description:	**All processing minicomputer-resident; all files maintained in central disk storage unit; testing stations function as input/output units only**
Processing Power/Cost:	**High power — high cost**
Resource Contention:	**Possible, especially in accessing CPU and disk**
System Availability:	**Direct relationship between number of testing stations and response degradation**
System Reliability:	**Dependent on central minicomputer and disk unit**
System Security:	**Hardware and software techniques applicable**
Portability:	**Not easily portable**
Operator Sophistication:	**Moderate**

Figure 18b. **Microcomputer-Based Configuration**

Technology:	**Sophisticated application of new technology**
Functional Description:	**Testing stations self-contained; functionally independent; Monitor station concentrates data and maintains network control**
Processing Power/Cost:	**High power — low to moderate cost**
Resource Contention:	**None**
System Availability:	**No relationship between number of testing stations and response degradation**
System Reliability:	**Dependent on number of testing stations likely to fail simultaneously; testing station failure does not crash system**
System Security:	**Hardware and software techniques applicable**
Portability:	**Easily portable**
Operator Sophistication:	**Minimal**

Figure 18. Comparison of minicomputer- and microcomputer-based CAT site hardware configurations.

The cost of using telecommunications to support a nationwide network of testing stations quickly becomes prohibitive (Civil Service Commission, 1979). One way to overcome the cost might be to install a minicomputer and supporting hardware at each site, with terminals serving as the monitoring and testing stations. As depicted in Figure 18a, this solution represents a straightforward application of established technology. All processing is minicomputer-resident, all files are maintained in a central disk storage unit, and the testing stations need to function only as input and output units. With the advent of 16-bit microprocessors, however, a microcomputer-based hardware configuration offers a promising alternative to the traditional minicomputer.

The microcomputer-based configuration (Figure 18b) represents a sophisticated application of new technology. Testing stations are self-contained, functionally independent units, each consisting of a microcomputer, disk unit, keyboard, and display. The monitor station is also self-contained; it serves to concentrate data from the testing stations and maintain control of the loosely coupled microcomputer network.

How do these configurations compare? The minicomputer offers high power at high cost, although the cost is much lower than that of a telecommunications network. The microcomputer also offers high power, at a lower cost than the minicomputer. In many other ways, microcomputers are preferable. Contention for resources is possible in the minicomputer configuration, especially in accessing the CPU and disk, while it is virtually nonexistent in the microcomputer configuration. In terms of system availability, the number of testing stations is directly related to the degree of response degradation in the minicomputer configuration. In terms of system reliability, failure of the minicomputer or its disk unit will crash the system and terminate all testing, while failure of a microcomputer-based testing station will only affect testing in progress at that station. For both configurations, current hardware and software security techniques would be applicable. For mobile site testing, the minicomputer configuration is not easily portable, while the microcomputer configuration provides easy portability. Finally, the minicomputer configuration normally requires moderate operator sophistication, while the microcomputer configuration requires minimal operator sophistication.

These comparisons are by no means definitive. They have been offered to suggest to systems designers that microcomputer technology shou'd be seriously considered in choosing the hardware configuration for CAT system testing sites. The performance characteristics of the new 16-bit microprocessors are impressive. Zilog (1978) claims that its Z8000 will outperform the Digital Equipment Corporation's PDP 11/45, a mid-range minicomputer. A recent article (Flippin, 1980) reports benchmark performance on a 16-bit multiply of 11 microseconds (μsec) for a Motorola 68000 microprocessor, compared with 10 μsec for an IBM 370-145, and 19 and 20 μsec respectively, for 2 other new 16-bit microprocessors, the Intel 8086 and the Zilog Z8000. This kind of performance should not be ignored. Although the system designer will probably have to configure a microcomputer-based system from the microprocessor up, so the speak, it may well be worth the effort. Characteristics of several selected minicomputers and microprocessors are provided in the appendix.

Software

The structural system design presented earlier in this report outlines the software requirements for the CAT system. Because this system software is primarily of the scientific, number-crunching type, FORTRAN, Pascal, or another high-level, structured programming language should be chosen for software development. Also, the complexity of the software design problem suggests that one of the structured software development techniques should be applied to ensure proper interfacing, protect system integrity, and aid in system documentation. Quality control of the software development effort is especially important, because the system's psychometric integrity is critically dependent on the degree to which system software accurately implements psychometric procedures.

Interfaces

Internal system interfaces have been discussed in the section on structural system design and are implied by the functional design model. Interface protocols will depend on the exact hardware configuration selected for the system. It should be noted, however, that interface design must reflect the data, the control paths, and the requirements specified in the functional model and structural design to assure smooth functioning of all components as an integrated system. The data and control requirements implied by the external interface to the MEPS reporting system must be carefully explored to ensure that the CAT system is successfully integreated with the enlisted personnel accessioning system.

Personnel

If the CAT system is to be successful, it must operate within the current accessioning environment and with present personnel. Both examinees and operating personnel must be considered. For examinees, the system must be "user friendly." Test-taking on the system must be simple and must present no threat. Software must be as forgiving of operating error as possible. Instructions must be clear and easily understood. The physical system must be human engineered for test-taking convenience. These requirements are also important for operating personnel; the system should be as fully automated as possible. Neither examinees nor operating personnel should be expected to have any degree of sophistication with regard to this type of system.

CAT System Testing, Evaluation, and Refinement

After the preliminary system design, the design's internal consistency and its external performance characteristics must be evaluated. Essentially, this involves verification of the design's logical consistency as it evolves from step to step, as well as validation of its ability to function according to specific system requirements (Enos & Van Tilburg, 1979). Verification and validation are carried out with regard to both function and structure. Performance evaluation seeks to determine performance characteristics that result from algorithmic design, system functional allocation and configuration, and structural interfaces. Computer simulation of the system processes that are amenable to such simulation (e.g., software module performance), as well as evaluation of system prototypes, the physical models of the system, provide necessary feedback on design decisions. Where applicable, computer simulation and prototype evaluation results are compared to check actual performance against the predicted performance of the system.[3]

The design testing, evaluation, and refinement step provides the last opportunity to make changes before full-scale implementation of the system design begins. This step must be carried out carefully and should meet applicable military standards (e.g., Military Standard: Technical Reviews and Audits for Systems, Equipment, and Computer Programs, MIL-STD-1521A, DoD, 1976).

---

[3] Colella, O'Sullivan, & Carlino (1974) have provided an excellent discussion of the rationale and precedures for system simulation and prototyping.

Functional Verification and Validation

Functional verification and validation refers to assurance that the functional design of the system is logically consistent and meets stated system objectives and requirements. This process answers the question of whether the system will do what it is supposed to do.

The process is applied to both psychometric and engineering development activities. In psychometric development, it ensures that the necessary processes implied by measurement theory have been well specified and integrated into an effective measurement system. For the CAT system design, it is necessary to understand thoroughly the system's theoretical base and its measurement algorithms, as well as the psychometric requirements and objectives of the design effort.

In engineering development, the process ensures that (1) the system's inputs, processes, and outputs have been specified in sufficient detail and in such a manner as to allow easy translation of function into the structure, logic, and organization of the system software, and (2) these functional specifications provide sufficient information to facilitate choices. For the CAT system design, it is necessary to understand software and hardware development and to appreciate the nature of the psychometric procedures to be implemented.

To be complete, verification and validation of the CAT system functional design must integrate psychometric and engineering concerns. A useful technique for functional verification and validation is the "structural walk-through," in which the design team meets to review the functional design, component by component, with an eye toward its internal consistency and the system objectives and requirements. This technique is especially useful for complex functional designs such as that of the CAT system. It should not be performed before the system's structural design is developed.

Structural Verification and Validation

Structural verification and validation refers to assurance that the structural design of the system is logically consistent and is an accurate translation of the functional design. This process answers the question of whether the system will perform its stated functions properly. Furthermore, it is a means of assuring that all system components fit into a well integrated whole. For systems such as CAT, in which functions are primarily implemented in software, structural verification and validation are oriented towards software testing and evaluation. Structured walk-throughs of organization, logic, and resultant program code will verify the accurate translation of the functional design into software. Simulation testing of the software at three levels (individual components, components integrated into individual subsystems, and subsystems integrated into full system design) serves as necessary validation of proper system functioning.

The design of the hardware configuration in which the system software will be implemented must also be subjected to this process. Especially in microprocessor-based configurations, where fairly low-level (e.g., chip or board) components must be effectively integrated, structural verification and validation provide the design checks necessary before funds are expended in prototype fabrication. Structured walk-throughs of hardware logic and organization, interfaces, and operating characteristics (processor speed, storage capacity and access time, and communication rates) verify the internal consistency of the design and validate expected performance characteristics of the hardware configuration. Simulation of system operation, staged either on partial prototype or the full system prototype, will confirm proper hardware and software functioning within the prototype-specific hardware context.

Structural verification and validation should be an integral part of the prototype development. This process is a necessary precursor to evaluation of the prototype in the performance evaluation phase and should be performed before prototyping of the system begins.

Performance Evaluation

Performance evaluation refers to assurance that the system will meet stated performance objectives in actual operation. It is primarily oriented towards prototype evaluation, through the application of simulation protocols that emulate real-world operating conditions. Developing those simulation protocols and the performance measures to be used in prototype evaluation is critical in evaluation of the system. The validity of the performance evaluation process will depend on the care taken in this development. Because the prototype represents a physical model of the system as it will operate in the real world, computer simulation will not suffice to test the prototype against all operating conditions. If the system is designed to test people and to be operated by people, the prototype must do so as well. Only when the prototype evaluation process represents a reasonable analog of real-world conditions will performance evaluation of the system be carried out successfully.

To assure that performance evaluation results will be meaningful, two prior conditions are important. First, evaluation criteria must be clearly and carefully specified, providing the metrics for comprehensive evaluation of system functioning against design objectives. Second, performance benchmarks for the evaluation criteria must be established, specifying the performance levels at which the prototype will be considered to have met or exceeded design objectives. These criteria and benchmarks must be established for both the psychometric and engineering aspects of the system design.

RECOMMENDATIONS

1. The design of the CAT system should be based on the 4 major functions and 25 subfunctions described in this report.

2. The HIPO technique, which is well suited to the problem of systematic top-down analysis of functional requirements, should continue to be employed throughout the evolution of the final CAT system design.

3. Although the CAT system could conceivably be based on a mainframe computer with a wide area network of remote terminals, telecommunication costs for such a system would be prohibitive. As alternatives, both microprocessors and minicomputers should be evaluated for their capabilities to support CAT test administration and the station-monitoring functions.

4. The 34 software components (subsystems, programs, modules, and subroutines) identified should serve as the basis for CAT system software development.

5. CAT's basis in mathematical statistics makes its implementation heavily dependent on scientific arithmetic computations; to support this requirement, FORTRAN, Pascal, or a similar high-level programming language should be used. Furthermore, the complexity of the CAT system functions and subfunctions suggests that structured software development techniques should be employed to facilitate software development, to protect system integrity, to ensure proper interfacing, and to aid in system documentation.

6. If the CAT system is to be cost-effective, it must be able to be operated by the user with operations staffs no larger than those required by the current system. Accordingly, one objective during CAT system design should be to minimize the number and skill requirements of personnel needed to operate and maintain the system.

7. The CAT system must meet stated system design objectives and requirements, from both hardware and software points of view. Meeting these objectives is best accomplished by means of a systematic process of testing, evaluation, and refinement. Formal procedures for design testing, evaluation, and refinement should be specified and used in the CAT system development process.

REFERENCES

Ackoff, R. L. Towards a system of systems concepts. In D. J. Cougar & R. W. Knapp (Eds.), Systems analysis techniques. New York: John Wiley & Sons, 1974.

Anastasi, A. Psychological testing (4th ed.). New York: Macmillan, 1976.

Bingham, J. E., & Davies, G. W. P. A handbook of systems analysis. London: The MacMillan Press, 1972.

Chapanis, A. Human factors in system engineering. In K. B. de Greene (Ed.). Systems psychology. New York: McGraw-Hill, 1970. (a)

Chapanis, A. System staffing. In K. B. de Greene (Ed.). Systems psychology. New York: McGraw-Hill, 1970. (b)

Civil Service Commission. First report of the computerized tailored testing study group. Washington, DC: Author, 1979. Internal document.

Clark, C. L. (Ed.). Proceedings of the First Conference on Computerized Adaptive Testing (PS-75-6, Personnel Research and Development Center, Civil Service Commission). Washington, DC: Government Printing Office, 1976. (Superintendent of Documents Stock No. 006-000-00940-9)

Colella, A. M., O'Sullivan, M. J., & Carlino, D. J. Systems simulation. Lexington, MA: D. C. Heath, 1974.

Croll, P. R., & Urry, V. W. Tailored testing: Maximizing validity and utility for job selection. Paper presented to Symposium on Innovations in Personnel Testing, 86th Annual Convention of the American Psychological Association, Toronto, 1975.

Datapro Research Corporation. Datapro reports on minicomputers (Vol. 1). Delran, NJ: Author, 1980.

de Greene, K. B. Systems Psychology. New York: McGraw-Hill, 1970.

Department of Defense. Military Standard: Technical reviews and audits for systems, equipment, and computer programs (MIL-STD-1521A). Washington, DC: Author, 1976.

Enos, J. C., & Van Tilburg, R. L. Software design. In R. W. Jensen & C. C. Tonies (Eds.), Software engineering. Englewood Cliffs, NJ: Prentice-Hall, 1979.

Flippin, A. The 16-bit time trials. Kilobaud Microcomputing, October 1980, 182-190.

IBM Corporation. HIPO--A design aid and documentation technique (2nd ed.). White Plains, NY: Author, 1975. (IBM Order No. GC20-1851-1)

Holtzman, W. H. (Ed.). Computer-assisted instruction, testing, and guidance. New York: Harper & Row, 1970.

Katzan, H. Systems design and documentation. An introduction to the HIPO method. New York: Van Nostrand Reinhold, 1976.

47

Lord, F. M. A broad-range test of verbal ability. Applied Psychological Measurement, 1977, 1, 95-100. (a)

Lord, F. M. Practical applications of item characteristic curve theory. Journal of Educational Measurement, 1977, 14, 177-128. (b)

McBride, J. R. Adaptive mental testing: The state of the art (Tech. Rep. 423). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences, November 1979.

McBride, J. R. Adaptive verbal ability testing in a military setting. Proceedings of the 1979 Computerized Adaptive Testing Conference. Minneapolis, MN: Psychometric Methods Program, University of Minnesota, September 1980.

McBride, J. R. Computerized adaptive testing project: Objectives and requirements (NPRDC TN 82-22). San Diego: Navy Personnel Research and Development Center, 1982.

Motorola Semiconductor Products. MC68000 advance information. Austin, TX: Author, 1979.

Ree, M. J., & Jensen, H. E. Effects of sample size on linear equating of item characteristic curve parameters. Proceedings of the 1979 Computerized Adaptive Testing Conference. Minneapolis, MN: Psychometric Methods Program, University of Minnesota, September 1980.

Rubin, M. L. Handbook of data processing management: Vol. 1, introduction to the system life cycle. Princeton, NJ: Brandon/Systems Press, 1970.

Schmidt, F. L., & Urry, V. W. Item parameterization procedures for the future. Proceedings of the First Conference on Computerized Adaptive Testing (PS-75-6, Personnel Research and Development Center, Civil Service Commission). Washington, DC: Government Printing Office, 1976. (Superintendent of Documents Stock No. 006-00-00940-9)

Urry, V. W. A Monte Carlo investigation of logistic mental test models. (Doctoral dissertation, Purdue University, 1970). Dissertation Abstracts International, 1971, 31, 6319B University Microfilms No. 71-9475.

Urry, V. W. Computerized-assisted testing: Calibration and evaluation of the verbal ability bank (TS-74-3). Washington, DC: Personnel Research and Development Center, Civil Service Commission, December 1974. (NTIS No. PB 261 694/AS)

Urry, V. W. Ancillary estimators for the item parameters of mental test models. In W. A. Gorham (Chair), Computers and testing: Steps toward the inevitable conquest. Symposium presented at the 83rd Annual Convention of the American Psychological Association, Chicago, 1975. (NTIS No. PB 261 694/AS)

Urry, V. W. A five-year quest: Is computerized adaptive testing feasible? Proceedings of the First Conference on Computerized Adaptive Testing (PS-75-6, Personnel Research and Development Center, Civil Service Commission). Washington, DC: Government Printing Office, 1976. (Superintendent of Documents Stock No. 006-00-00940-9)

Urry, V. W. Tailored testing: A successful application of latent trait theory. Journal of Education Measurement, 1977, 14, 181-196. (a)

Urry, V. W. Tailored testing: A spectacular success for latent trait theory (TS-77-2). Washington, DC: Personnel Research and Development Center, Civil Service Commission, August 1977. (NTIS No. PB 274-576) (b)

Urry, V. W. A multivariate model sampling procedure and method of multidimensional tailored testing. Proceedings of the 1977 Computerized Testing Conference. Minneapolis, MN: Psychometric Methods Program, University of Minnesota, 1980.

Urry, V. W. Tailored testing: Its theory and practice. Part 1. The basic model, the normal ogive submodels, and the tailored testing algorithms. Manuscript submitted for publication, 1981. (a)

Urry, V. W. Tailored testing: Its theory and practice. Part 2. Ability and item parameter estimation, multiple ability applications, and allied procedures. Manuscript submitted for publication, 1981. (b)

Weiss, D. J. (Ed.). Proceedings of the 1977 Computerized Adaptive Testing Conference. Minneapolis, MN: Psychometric Methods Program, University of Minnesota, 1978.

Weiss, D. J. (Ed.). Proceedings of the 1979 Computerized Adaptive Testing Conference. Minneapolis, MN: Psychometric Methods Program, University of Minnesota, 1980.

Zilog, Inc. Zilog Z800: An architectural overview. Cupertino, CA: Author, 1978.

APPENDIX

CHARACTERISTICS OF SELECTED DATA PROCESSING HARDWARE

This appendix lists specifications for eight minicomputers and eight microprocessors that represent the range of equipment available in the current market. The selections have concentrated on 16-bit machines because their high performance makes them more suitable than the 8-bit machines for the heavy number-crunching tasks in computerized adaptive test administration and scoring.

It should be noted that, for all the microprocessors listed, compatible parts are available that allow them to be incorporated into a microcomputer design (e.g., random-access memory, read-only memory, input/output interfaces, clock generators). These processors must be incorporated into such a design to support computerized adaptive test administration and scoring.

Except for the information on the MC 6800, which was excerpted from vendor literature (Motorola, 1979), the information presented herein was excerpted from the Datapro Reports on Minicomputers, Volume 1 (Datapro, 1980) and used with permission.

Table A-1

Characteristics of Selected Data Processing Hardware

Characteristics	Minicomputer							
	Data General Eclipse S/140	Data General Nova 4X	Digital Equipment PDP-11/70	Digital Equipment PDP-11/70	Hewlett-Packard HP1000F	Modular Computer Systems, Inc. Classic 7830/7835	Harris 100	Systems Engineering Laboratories 32/77
Word length, bits	16 + 5	16	16 + 2	16 + 2	16 + 1	16	24, 48	32 + 7
Number of terminals supported	64	64	—	—	56	96	Application Dependent	64
MAIN STORAGE								
Cycle access time, microseconds	0.20 0.40	—	0.48, 0.96/0.48	0.98/0.36	0.35	.125/.250	0.45/0.30	0.60/0.30
Min./max. capacity words	64K 512K	64K/128K	256K/1M bytes	64K/1024K	32K/2048K bytes	128K/2048K bytes	192K/768K bytes	64K/4096K
Parity checking	No	No	No	Standard	Standard	Standard	No	No
Error correction	Standard	No	Standard	No	Optional	Standard	Standard	Standard
Storage protection	Standard	No	Standard	Standard	Optional	Standard	Standard	Standard
CENTRAL PROCESSOR								
Number of directly addressable words	32K	1K	32K	32K	2K	2048K	96K	128K
Add time, microseconds	0.20	0.20	0.87	0.30-1.20	0.91	0.30	0.60	0.60/1.20
Hardware multiply divide	Standard	Optional	Standard	Standard	Standard	Standard	Standard	Standard
Hardware floating point	Optional	Optional	Optional	Optional	Firmware	Optional Standard	Optional	Standard
Hardware byte manipulation	Standard	Standard	Standard	Standard	Standard	Standard	Standard	Standard
Real time clock or timer	Standard	Standard	Standard	Standard	Optional	Standard	Optional	Standard
Direct memory access	Standard	Standard	Standard	Standard	Optional	Standard	Optional	Standard
COMMUNICATIONS								
Maximum number of lines	—	128	—	—	56	256 FOX	32	64
Synchronous	Opt.; 56K bps	Opt.; (32) 56K bps	Up to 1M bps	Up to 1M bps	Opt.; to 500K bps	Opt.; 48-230.4K bps	Opt.; 56K bps	Opt.; to 9600 bps
Asynchronous	Opt., 9600 bps	Opt.; (128) 19.2K bps	Up to 9600 bps	Up to 9600 bps	Opt.; to 2.5M bps	Opt.; 50-19.2K bps	Opt.; 19.2K bps	Opt.; to 38.4 bps
Higher level language(s)	BASIC, FORTRAN	BASIC, FORTRAN	BASIC, FORTRAN	BASIC, FORTRAN	FORTRAN, BASIC	FORTRAN	FORTRAN IV & 77	FORTRAN, BASIC
Operating system	Batch, real-time, time-sharing	Real-time, RDOS, multi-tasking	Batch, real-time, time-sharing	Real-time, interactive, time-sharing	Real-time, time-sharing	Batch, real-time, time-sharing	Real-time, batch, time-sharing	Real-time, interactive, multi-batch
Price of CPU, power supply, front panel, and minimum memory in chassis	$16,500 (128K bytes)	$10,400 (128K bytes)	$23,900 (256K bytes)	$63,000 (128K core)	$11,750 (64K bytes)	$23,800/29,500	$45,000 (192K bytes)	$46,300 (256K bytes)

Table A-1 (Continued)

Characteristics	Microprocessor							
	Intel 8085A	Intel 8086-2	Intel 8087	Intel 8089	Motorola 6800	Motorola 68000	Zilog Z80A/Z80B	Zilog Z8001
Type	8-bit CPU	16-bit CPU	—	8/16-bit I/O processor	8-bit CPU	16-bit CPU	8-bit CPU	16-bit CPU
Data word size, bits	8, 16, 24	8-48	16, 32, 64, 80	8-16	8	16 (varies, 1-32 bits)	8	16 (varies, 1-32 bits)
Instruction word size, bits	8, 16, 24	8-48	16-48	16	8, 16, 24	16-80	8, 16	16
Clock frequency	3, 5 MHz	5 MHz	5 MHz	5 MHz	1 MHz	To 8 MHz	2.5, 4.0, or 6.0 MHz	To 6 MHz
Phases/cycle	1	1	1	1	2	1	1	1
Add time, register to register, microseconds per data word	1.0	0.6 (8 or 16-bit)	0.2 (64-bit add)	—	2.0	0.5	1.6	1.0
Number of instructions	82	134	58	45	72	56	158	110
NUMBER OF REGISTERS:								
Arithmetic	1	—	—	—	Two 8-bit	8 32-bit	14	—
Index	0	—	—	—	One 16-bit	Up to 17	Two 16-bit	—
General purpose	6	8 8- or 16-bit; 4 memory segmentation	8x8 bit	8 20-bit 8 16-bit	Two 8-bit	7 32-bit	Two sets of six each	16
Size of return stack	Unlimited	Unlimited	Unlimited	—	Up to 64K	Unlimited	Unlimited	—
Number of directly addressable inctruction words	64K	1M	1M	1M + 64K	64K	16M	64K	8M
Hardware BCD arithmetic	No	Standard	Yes	No	Standard	Standard	Standard	Standard
Direct memory access	Optional	Optional	—	—	Available	Standard	Standard	Standard
Higher level languages	PLM-80, PASCAL, BASIC	PLM-86	PLM-86, FORTRAN, PASCAL	No	MPL, BASIC	PASCAL	PL/z, FORTRAN, PASCAL	PASCAL
Price of basic CPU only (quantity 100)	$11.25	$112.50	Contact vendor	Contact vendor	$13.75 (25-99)	Contact vendor	$8.90/$10.70	$140
Comments		8 and 16-bit signed/unsigned arithmetic, including multiply and divide	Ultra high performance numeric data co-processor for 8086	I/O co-processor for 8086				Segmented version of CPU. Specifications taken from second-sourced advanced Micro Devices AM Z8001

DISTRIBUTION LIST

Director Accession Policy (OASD(Military Personnel and Force Management)) (2)
Military Assistant for Training and Personnel Technology (ODUSD(R&AT)) (2)
Assistant Secretary of the Navy (Manpower and Reserve Affairs) (2)
Deputy Assistant Secretary of the Navy (Manpower) (OASN(M&RA)) (2)
Director of Manpower Analysis (ODASN(M))
Chief of Naval Operations (OP-01), (OP-11), (OP-12) (2), (OP-13), (OP-14), (OP-15), (OP-102), (OP-115) (2), (OP-135), (OP-140F2), (OP-964D), (OP-987H)
Chief of Naval Material (NMAT 00), (NMAT 05) (2), (NMAT 08)
Chief of Naval Research (Code 200), (Code 440) (3), (Code 442), (Code 448), (Code 450) (6)
Chief of Information (OI-213)
Chief of Naval Education and Training (02), (015), (N-5)
Chief of Naval Technical Training (016)
Commandant of the Marine Corps (MPI-20) (12)
Commander Fleet Training Group,, Pearl Harbor
Commander Naval Data Automation Command
Commander Naval Military Personnel Command (NMPC-013C)
Commander Navy Recruiting Command
Commanding Officer, Naval Aerospace Medical Institute (Library Code 12) (2)
Commanding Officer, Naval Regional Medical Center, Portsmouth, VA (ATTN: Medical Library)
Director, Naval Civilian Personnel Command
Deputy Chief of Staff for Personnel (DAPE-MPE)
Commander, Army Research Institute for the Behavioral and Social Sciences, Alexandria (PERI-ASL), (PERI-ZT) (6)
Headquarters Commandant, Military Enlistment Processing Command, Fort Sheridan (6)
Chief, Army Research Institute Field Unit, Fort Harrison
Commander, Air Force Human Resources Laboratory, Brooks Air Force Base (Scientific and Technical Information Office)
Commander, Air Force Human Resources Laboratory, Brooks Air Force Base (Code MPA) (2)
Commander, Air Force Human Resources Laboratory, Lowry Air Force Base (Technical Training Branch) (2)
Commander, Air Force Human Resources Laboratory, Williams Air Force Base (AFHRL/OT)
Commander, Air Force Human Resources Laboratory, Wright-Patterson Air Force Base
Commander, Air Force Military Personnel Command, Randolph Air Force Base (Code MPCYPT)
Program Manager, Life Sciences Directorate, Bolling Air Forcce Base (2)
Commanding Officer, U.S. Coast Guard Institute (2)
Superintendent, U.S. Coast Guard Academy
Defense Technical Information Center (DDA) (12)