

THE DEVELOPMENT AND APPLICATION OF MODELS TO PREDICT OPERATOR WORKLOAD  
DURING SYSTEM DESIGN\*

Theodore B. Aldrich

Anacapa Sciences Inc.  
Fort Rucker, Alabama

Sandra M. Szabo

Anacapa Sciences Inc.  
St. Louis, Missouri

Carl R. Bierbaum

Anacapa Sciences Inc.  
Fort Rucker, Alabama

INTRODUCTION

New weapons being developed for modern military forces feature advanced technology designed to extend and improve mission performance beyond the capability of existing systems. For example, aircraft systems are being developed with advanced technology designed to extend range, increase speed, provide for more precise navigation, avoid enemy threats, and acquire and engage enemy targets at night or in adverse weather.

In addition to improving mission capability, advanced technology also is designed to reduce crew workload. However, in some instances the tasks required to operate the technology may actually increase workload. The increased workload, in turn, may degrade human performance and, consequently, reduce rather than improve mission effectiveness.

Models that predict operator workload can be useful tools for human factors engineers who are attempting to address human capabilities and limitations as advanced technology is introduced into new weapon systems. In response to this requirement Anacapa Sciences, Inc. researchers, under contract to the U.S. Army Research Institute Aviation Research and Development Activity, have developed a series of models for predicting aviator workload. The work supports U.S. Army design studies for the following helicopter systems:

- a highly automated, multipurpose, lightweight helicopter, designated LHX (see Aldrich, Szabo, & Craddock, 1986);

---

\*Research reported in this chapter was performed in support of the Army Research Institute Aviation Research and Development Activity, Fort Rucker, Alabama, under Contract Numbers MDA903-81-C-0504 and MDA903-87-C-0523. Mr. Charles A. Gainer was the technical monitor. The views expressed are those of the authors and are not necessarily endorsed by the U.S. Army.

- the AH-64A, Apache (see Szabo & Bierbaum, 1986);
- the UH-60A, Blackhawk (see Bierbaum, Szabo, & Aldrich, 1987); and
- the CH-47D, Chinook.

The LHX models were used in advance of system design to predict single- and dual-crew workload under varying automation configurations. The AH-64A, UH-60A, and CH-47D models presently are being used for evaluating the impact of advanced technology modifications being proposed for each of these existing helicopters.

This chapter describes a four-phase research program aimed at the development and application of models to predict operator workload during system design. Phase 1 consists of the development of a mission/task/workload analysis data base. Phase 2 consists of the development of computer models to predict operator workload. Phase 3 consists of applying the workload prediction models during system design studies. Phase 4 consists of research required to validate the workload predictions yielded by the models. Most of this chapter describes research performed in support of the LHX as reported by Aldrich, Szabo, and Craddock (1986). However, refinements in the methodology introduced by Szabo and Bierbaum (1986) in support of proposed AH-64A modifications, and by Bierbaum, Szabo, and Aldrich (1987) in support of proposed UH-60A modifications, also are included. Thus, this chapter presents the current state of the Anacapa Sciences, Inc. research directed at developing and validating operator workload prediction models.

#### THE MISSION/TASK/WORKLOAD ANALYSIS DATA BASE

The first phase of the workload prediction methodology requires the conduct of a comprehensive mission/task/workload analysis. In the case of the LHX, 24 proposed scout and attack mission profiles provided by the Directorate of Combat Developments (DCD) at the U.S Army Aviation Center (USAAVNC), Fort Rucker, Alabama were examined. Because of program schedule constraints only nine of the 24 mission profiles were selected for preliminary analysis. The nine mission profiles were subsequently divided into mission phases; the following three mission phases were selected for detailed analysis:

- Reconnaissance,
- Target Service (Air-To-Ground), and
- Target Service (Air-To-Air).

Each of the three mission phases listed above was further divided into segments; a limited sample of 29 mission segments was selected for the detailed task analysis. Each of the 29 mission segments, in turn, was divided into mission functions. Finally, each of the mission functions was divided into mission tasks. A total of 58 unique functions and 135 unique tasks were identified for the 29 mission segments that were analyzed.

The same general procedure was used to conduct the AH-64A, UH-60A, and CH-47D analyses. In the case of the AH-64A, Szabo and Bierbaum (1986) conducted a comprehensive analysis of an entire composite mission from preflight through postflight. They identified 52 unique segments, 159 unique functions, and 689 unique tasks. In the UH-60A analysis, Bierbaum, Szabo, and Aldrich (1987) identified 34 critical segments, which were subsequently divided into 48 unique functions and 138 unique tasks. In the CH-47D analysis, 37 critical mission segments were divided into 65 unique functions and 154 unique tasks.

For each unique task, the following additional data were derived:

- crewmember performing the task,
- subsystem equipment associated with the task,
- estimate of the time required to perform the task, and

- estimates of workload associated with the sensory (i.e., visual, auditory, kinesthetic) cognitive, and psychomotor components of the task.<sup>1</sup>

Procedures for deriving the additional task data are briefly described in the paragraphs that follow.

#### Crewmember Performing the Task

The first step in deriving the additional task data was to identify the crewmember performing each task. All flight control tasks were assigned to the pilot. Primary mission tasks (e.g., Align Sight, Activate Trigger) and most support tasks (e.g., Check Aircraft Systems, Transmit Message) were assigned to the second crewmember.

#### Subsystem Equipment Associated With the Task

The next step in the analyses was to identify the subsystem equipment associated with the performance of each task. In each analysis the identified subsystem equipments were categorized into major subsystem categories. The categories vary among the different systems depending upon (a) the mission assigned to the particular aircraft of interest and (b) the existing configuration of that aircraft.

#### Estimate of the Time Required to Perform the Task

The methods of estimating task times also varied somewhat for the different systems. Aldrich, Craddock, and McCracken (1984) describe the methods for estimating task times in the LHX analyses. In their analyses, each task was first categorized as discrete or continuous. Discrete tasks are characterized by actions having a definite, observable start and end point. Activation of switches, performance of procedures, and transmissions of radio messages are examples of discrete tasks. Existing helicopter task analyses for the OH-58D (Taylor & Poole, 1983) and for the AH-64 and the Advanced Helicopter Improvement Program (AHIP) (Siegal, Madden, & Pfeiffer, 1985) were used as references in deriving estimates of LHX discrete task times.

Continuous tasks do not have observable start and end points and cannot be reduced to procedures; mission requirements and conditions determine their duration. Examples of continuous tasks are flight control tasks and target tracking tasks. Aldrich, Craddock, and McCracken (1984) assigned times to continuous tasks so that each discrete task could be accomplished within the elapsed times assigned to concurrent continuous tasks. For example, the times assigned to the continuous tasks associated with the Hover Masked function, were long enough to allow the operator to complete all of the discrete tasks (e.g., Check Aircraft Systems, Transmit Message) performed concurrently with the continuous tasks in the Hover Masked function. All assigned times for discrete and continuous tasks were reviewed by subject matter experts (SMEs).

During the AH-64A analysis, Szabo and Bierbaum (1986) identified two types of discrete tasks. Specifically, they categorized discrete tasks as either "discrete fixed" or "discrete random". Discrete fixed tasks have definite start and end points within the function (e.g., Set SIGHT SEL Switch). Discrete random tasks are discrete tasks that occur intermittently and/or randomly during a portion of the function (e.g.,

<sup>1</sup>Estimates of the kinesthetic workload component of tasks were introduced during the AH-64A analysis by Szabo and Bierbaum (1986). The higher specificity of their task analysis, compared to the LHX analyses, required the kinesthetic estimates. The kinesthetic estimates were retained by Bierbaum, Szabo, and Aldrich (1987) in their UH-60A analysis and are currently being used in the CH-47D analysis.

Check Fuel Quantity Indicator). Szabo and Bierbaum derived most of their task times by timing the actual tasks as they were performed in the AH-64A Cockpit, Weapons, and Emergency Procedures Trainer. For tasks not trainable in the trainer they used estimates provided by AH-64A SME's. Bierbaum, Szabo, and Aldrich (1987) retained the refined categorization of discrete tasks for the UH-60A analysis. UH-60A task time estimates were obtained during interviews with UH-60A SMEs.

#### Estimates of Workload Associated With the Sensory, Cognitive, and Psychomotor Components of the Task

Workload, as the term is used in this research, is defined as the total attentional demand (i.e., mental workload) placed on the operator(s) as they perform the mission tasks. Consistent with Wickens theory that workload is a multidimensional construct, the research methodology addresses three different components of workload; sensory, cognitive, and psychomotor (Wickens, 1984). The sensory component refers to the complexity of the visual (V), auditory (A), or kinesthetic (K) stimuli to which an operator must attend; the cognitive (C) component refers to the level of information processing required from the operator; the psychomotor (P) component refers to the complexity of the operator's behavioral responses. The steps performed to determine the workload associated with each of these components for each of the mission tasks are described in the paragraphs that follow.

McCracken and Aldrich (1984) estimated LHX task workload by using 7-point ordinal scales for rating the visual, cognitive, and psychomotor workload components and a 4-point ordinal scale for rating the auditory workload components of each task. Szabo and Bierbaum (1986) added a kinesthetic sensory component to their analysis of workload, and developed an ordinal 7-point kinesthetic rating scale with verbal anchors similar to the visual, cognitive, and psychomotor rating scales. They also developed an ordinal 7-point auditory rating scale to replace the original 4-point auditory rating scale used by McCracken and Aldrich.

During the UH-60A analysis, Bierbaum, Szabo, and Aldrich (1987) added a second visual scale and converted the ordinal scale measures to interval scale measures. The second visual scale was added so that the attentional demand associated with the visual component of the mission tasks could be estimated under both naked eye (visual-unaided) and night vision goggle (visual-aided) conditions. Both visual scales retain the same verbal anchors used in the prediction of AH-64A crew workload.

The interval scales used in the UH-60A analysis were constructed by using a pair comparison survey methodology (Engen, 1971). The survey presented matched pairs of verbal anchors for the visual (both naked eye and night vision goggles), auditory, cognitive, and psychomotor workload component scales to 20 UH-60A instructor pilots (IPs) from the UH-60A Aviator Qualification Course (AQC) at the USAAVNC, Fort Rucker, Alabama. The frequency with which the IPs selected each verbal anchor was used to compute a value for each verbal anchor on an approximately equal-interval scale.

The matched pairs of verbal anchors for the kinesthetic workload component scale were similarly arranged in a questionnaire and administered by mail to a group of 22 human factors experts who have had extensive research experience in workload measurement. Pair comparison response frequencies were tabulated to develop interval scale values for the kinesthetic workload component scale. The six workload component interval scales used in the UH-60A analysis are presented in Table 1.

Once the workload component scales had been developed, a short verbal descriptor of each of the workload components was written for each task. The descriptors were then compared to the verbal anchors in the appropriate interval or rating scale. In each instance, a consensus was

**Table. 1.** Workload Component Scales for the UH-60A Mission/Task/Workload Analysis

| Scale Value | Descriptors                                                                                |
|-------------|--------------------------------------------------------------------------------------------|
|             | Visual-Unaided (Naked Eye)                                                                 |
| 1.0         | Visually Register/Detect (Detect Occurrence of Image)                                      |
| 3.7         | Visually Discriminate (Detect Visual Differences)                                          |
| 4.0         | Visually Inspect/Check (Discrete Inspection/Static Condition)                              |
| 5.0         | Visually Locate/Align (Selective Orientation)                                              |
| 5.4         | Visually Track/Follow (Maintain Orientation)                                               |
| 5.9         | Visually Read (Symbol)                                                                     |
| 7.0         | Visually Scan/Search/Monitor (Continuous/Serial Inspection, Multiple Conditions)           |
|             | Visual-Aided (Night Vision Goggles [NVG])                                                  |
| 1.0         | Visually Register/Detect (Detect Occurrence of Image) With NVG                             |
| 4.8         | Visually Inspect/Check (Discrete Inspection/Static Condition (With NVG)                    |
| 5.0         | Visually Discriminate (Detect Visual Differences) With NVG                                 |
| 5.6         | Visually Locate/Align (Selective Orientation) With NVG                                     |
| 6.4         | Visually Track/Follow (Maintain Orientation) With NVG                                      |
| 7.0         | Visually Scan/Search/Monitor (Continuous/Serial Inspection, Multiple Conditions (With NVG) |
|             | Auditory                                                                                   |
| 1.0         | Detect/Register Sound (Detect Occurrence of Sound)                                         |
| 2.0         | Orient to Sound (General Orientation/Attention)                                            |
| 4.2         | Orient to Sound (Selective Orientation/Attention)                                          |
| 4.3         | Verify Auditory Feedback (Detect Occurrence of Anticipated Sound)                          |
| 4.9         | Interpret Semantic Content (Speech)                                                        |
| 6.6         | Discriminate Sound Characteristics (Detect Auditory Differences)                           |
| 7.0         | Interpret Sound Patterns (Pulse Rates, Etc.)                                               |
|             | Kinesthetic                                                                                |
| 1.0         | Detect Discrete Activation of Switch (Toggle, Trigger, Button)                             |
| 4.0         | Detect Preset Position or Status of Object                                                 |
| 4.8         | Detect Discrete Adjustment of Switch (Discrete Rotary or Discrete Lever Position)          |
| 5.5         | Detect Serial Movements (Keyboard Entries)                                                 |
| 6.1         | Detect Kinesthetic Cues Conflicting with Visual Cues                                       |
| 6.7         | Detect Continuous Adjustment of Switches (Rotary Rheostat, Thumbwheel)                     |
| 7.0         | Detect Continuous Adjustment of Controls                                                   |
|             | Cognitive                                                                                  |
| 1.0         | Automatic (Simple Association)                                                             |
| 1.2         | Alternative Selection                                                                      |
| 3.7         | Sign/Signal Recognition                                                                    |
| 4.6         | Evaluation/Judgment (Consider Single Aspect)                                               |
| 5.3         | Encoding/Decoding, Recall                                                                  |
| 6.8         | Evaluation/Judgment (Consider Several Aspects)                                             |
| 7.0         | Estimation, Calculation, Conversion                                                        |

(continued)

Table. 1. Workload Component Scales for the UH-60A Mission/Task/Workload Analysis (Continued)

| Scale Value | Descriptors                                                      |
|-------------|------------------------------------------------------------------|
|             | Psychomotor                                                      |
| 1.0         | Speech                                                           |
| 2.2         | Discrete Actuation (Button, Toggle, Trigger)                     |
| 2.6         | Continuous Adjustive (Flight Control, Sensor Control)            |
| 4.6         | Manipulative                                                     |
| 5.8         | Discrete Adjustive (Rotary, Vertical Thumbwheel, Lever Position) |
| 6.5         | Symbolic Production (Writing)                                    |
| 7.0         | Serial Discrete Manipulation (Keyboard Entries)                  |

reached by the two analysts who initially had assigned the workload estimates independently. The consensual estimates were subsequently reviewed by SMEs for the selected system.

A complete summary of the data derived from the mission/task/workload analysis was entered on function analysis worksheets, such as the one selected from the AH-64A analysis (Szabo and Bierbaum, 1986) and depicted in Figure 1. A separate worksheet was prepared for each unique function identified in each analysis. The verb and object for each task within the function are presented in the first two columns, respectively. The crewmember performing each task is indicated by the letter (i.e. Pilot [P], Gunner [G] or Both [B]) in the third column. The subsystems associated with each task are shown in the fourth column. The verbal descriptors and the numerical estimates of workload for the sensory, cognitive, and psychomotor components (i.e., Visual-Unaided [V], Visual-Aided [G], Auditory, [A], Kinesthetic [K], Cognitive [C] and Psychomotor [P]) of each task are shown in the fifth, sixth, and seventh columns. For each task involving a specific switch, a switch description is presented in the eighth column. The estimated length of the discrete tasks is presented in the ninth column. The continuous tasks are identified in the tenth column with the letter "c." The function analysis worksheets thus provide a comprehensive summary of the information used to establish the data base for developing the workload prediction models in Phase 2 of the research.

FUNCTION 54 Designate Target (Autonomous) TOTAL TIME (Approximate) 13.5 Seconds

| TASKS   |                   | ID # | SUBSYSTEM(S)                | WORKLOAD COMPONENTS       |                                                                    |                                       | SWITCH DESCRIPTION          | DURATION (SECONDS) DISCRETE/ CONTINUOUS |  |
|---------|-------------------|------|-----------------------------|---------------------------|--------------------------------------------------------------------|---------------------------------------|-----------------------------|-----------------------------------------|--|
| VERB    | OBJECT            |      |                             | SENSORY                   | COGNITIVE                                                          | PSYCHOMOTOR                           |                             |                                         |  |
| Monitor | HAD Message (TOF) | G296 | Fire Control Computer (AFC) | Read Symbolic Display V-7 | Interpret Symbolic Readout and Make Judgment (Time to Lase) C-5    |                                       |                             | 5                                       |  |
| Pull    | Laser Trigger     | G361 | Laser (AL)                  | Feel Trigger Movement K-2 | Verify Correct Position (Laser Activated) C-2                      | Lift Cover, Pull and Hold Trigger P-1 | Springloaded Trigger (SPTR) | 1                                       |  |
| Note    | Weapon Impact     | G639 | Sensor Display (VSD)        | Visually Detect Image V-1 | Evaluate Sensory Feedback and Make Judgment (Target Destroyed) C-5 |                                       |                             | 5                                       |  |
| Release | Laser Trigger     | G362 | Laser (AL)                  | Feel Trigger Movement K-2 | S-R Association C-1                                                | Release Trigger P-1                   | Springloaded Trigger (SPTR) | .5                                      |  |

Figure 1. AH-64A function analysis worksheet.

## DEVELOPMENT OF COMPUTER BASED WORKLOAD PREDICTION MODELS

Phase 2 of the methodology consists of developing computer models to predict total workload experienced in the performance of individual and concurrent tasks. Whereas the mission/task/workload analysis methodology follows a top-down approach, the computer models are developed using a bottom-up approach. The task data identified during Phase 1 constitute the basic elements of analysis. The steps required to develop the models follow:

- establish computer data files,
- write function and segment decision rules, and
- write computer programs.

Each of these steps is described briefly in the subsections below.

### Establish Computer Data Files

The first step in developing each of the workload prediction models is to enter the mission/task/workload data derived during Phase 1 into computer files. Specifically, the information summarized on the function analysis worksheets is used to create the following data files:

- a list of segments,
- a list of functions,
- a list of tasks,
- a list of subsystem identifiers,
- workload ratings, and
- time estimates.

### Develop Function and Segment Decision Rules

The next step in developing the workload prediction models is to write time-based decision rules for building the mission segments from the task data base. Function decision rules specify the sequence and time for the performance of each task within each function; segment decision rules specify the sequence and temporal relationships for combining the functions to form mission segments. For the LHX analyses, Aldrich, Craddock, and McCracken (1984) developed one set of segment decision rules for a one-crewmember configuration and a second set of decision rules for the two-crewmember LHX configuration. Szabo and Bierbaum (1986) developed a single set of segment decision rules for the AH-64A analysis, and Bierbaum, Szabo and Aldrich (1987) developed another set of segment decision rules for the UH-60A analysis.

### Write Computer Programs

The time-based function and segment decision rules are the blueprints for placing the tasks performed by the operator(s) at the appropriate point on the mission timeline. Computer programs are developed to implement the function and segment decision rules. The timeline produced by programming the function and segment decision rules enables the identification of all tasks performed by the operator(s) at each half-second interval in the mission segment.

Computer programs also are developed for producing estimates of total workload associated with the performance of concurrent and sequential tasks. The total workload for concurrent tasks is computed by summing the workload component ratings (i.e., visual, auditory, kinesthetic, cognitive, and psychomotor) assigned during the task analyses. The specific half-second intervals when excessive workload occurs can be identified on the segment timeline by referring to the workload component sums. Four indices of overload producible by the model have been developed (Aldrich, Craddock, & McCracken 1984) and are listed and defined below:

- A component overload occurs whenever the sum of the ratings assigned to a given workload component (i.e., visual, auditory, kinesthetic, cognitive, or psychomotor) for concurrent tasks equals "8" or higher. Thus as many as five component overloads may occur for two or more concurrent tasks. A value of "8" was chosen as the criterion for an overload because it exceeds the maximum value on any of the workload component rating scales.
- An overload condition occurs whenever a component overload, as defined above, occurs in at least one component of the concurrent tasks. In theory as many as five component overloads (i.e., visual, auditory, kinesthetic, cognitive, and psychomotor) may occur within a single overload condition.
- Overload density is the percentage of time during a mission segment that a component overload occurs. It is calculated by dividing the number of timelines with component overloads by the total number of timelines in the segment.
- The term subsystem overload is used to describe the relationship between a component overload and a subsystem. It is computed by tallying the number of times each subsystem is associated with a component overload.

The component overload, overload condition, and subsystem overload indices provide diagnostic information about excessive workload for concurrent tasks. The overload density index provides a potential diagnostic measure of cumulative workload associated with sequences of concurrent tasks.

Following the steps described above, Aldrich, Craddock, and McCracken (1984) developed both one- and two-crewmember baseline workload prediction models for LHX analyses. Workload prediction models also have been developed for the AH-64A (Szabo & Bierbaum, 1986) and for the UH-60A (Bierbaum, Szabo, & Aldrich, 1987). These baseline workload prediction models provide benchmarks for comparisons to be made when the models are exercised to predict workload for alternative crew configurations or proposed automation options.

#### APPLYING THE WORKLOAD PREDICTION MODELS DURING SYSTEM DESIGN STUDIES

The third phase of the research consists of exercising the workload prediction models and applying the results to system design studies. This section describes the third phase of the research and presents some of the results produced from applying the LHX workload prediction models.

##### Workload Predictions: One- vs Two-Crewmember LHX Baseline Configurations

The one- and two-crewmember baseline LHX workload prediction models were developed using the data base compiled during the LHX mission/task/workload analysis. The tasks, subsystems, workload ratings, and time estimates are identical in both models and the function and segment decision rules were written so that both models have identical timelines. The only difference between the two models is the allocation of the functions between the crewmembers. Thus, workload predictions produced by the one-crewmember baseline model can be compared with workload predictions produced by the two-crewmember model to provide estimated differences in operator workload between the one- and two-crewmember LHX configurations.

Results summarized in Table 2 indicate that, for the 29 LHX segments, there were 263 overload conditions in the baseline one-crewmember configuration and 43 overload conditions in the baseline two-crewmember configuration. The 263 overload conditions in the



Table 2. Frequency of Overload Conditions and Component Overloads: One and Two-Crewmember LHX Baseline Configuration

|                | Number of<br>Overload Conditions | Number of<br>Component Overloads |    |    |     |       |
|----------------|----------------------------------|----------------------------------|----|----|-----|-------|
|                |                                  | V                                | A  | C  | P   | Total |
| One-Crewmember | 263                              | 79                               | -- | 54 | 203 | 336   |
| Two-Crewmember | 43                               | 21                               | -- | 17 | 15  | 53    |

one-crewmember configuration are composed of 79 visual component overloads, 54 cognitive component overloads, and 203 psychomotor component overloads, for a total of 336 component overloads. The 43 overload conditions in the two-crewmember configuration are composed of 21 visual component overloads, 17 cognitive component overloads, and 15 psychomotor component overloads, for a total of 53 component overloads. In the one-crewmember configuration overload conditions were predicted in each of the 29 segments that were analyzed. In the two-crewmember configuration, overload conditions were predicted in only 15 of the 29 segments; the pilot was overloaded in only three of these 15 segments.

#### Workload Predictions: One- vs Two-Crewmember Configurations With Proposed Automation Options

The next step in the LHX analyses was to exercise the one- and two-crewmember models to predict how much operator workload would be reduced by individual automation options and combinations of options being considered for the LHX design. The methodology consists of three tasks:

- selecting automation options to be exercised by the models,
- revising the estimates of workload for each task, and
- exercising the one- and two-crewmember computer models.

Selecting the automation options. As part of the Army's LHX trade-off studies, the DCD at the USAAVNC, Fort Rucker, Alabama, developed alternative mission equipment packages (MEP) and aircraft survivability equipment (ASE) packages for the LHX. The MEP and ASE consisted of advanced technology equipments designed to automate many of the crew functions. The MEP and ASE descriptions were reviewed by Anacapa analysts and human factors specialists assigned to the DCD. Twenty-six individual automation options of interest were selected for analysis.

Revising the workload estimates. The next step in applying the methodology was to determine how each of the automation options would affect operator workload. A review of the task descriptions and the generic subsystems reported on the function analysis worksheets provided clues about how the workload would be affected by each of the proposed automation options. Based on the review, new descriptors of the operator's activities were entered into the sensory, cognitive, and psychomotor columns of the worksheets. The revised descriptors were then used to assign new estimates of workload to each component of those tasks affected by the automation options. In cases where automation completely eliminated a task, zero ratings were assigned to the workload components. No time estimates were changed as a function of automation; therefore, the decision rules for building functions from tasks and for building segments from functions remained unchanged.

Exercising the models with the automation options. Following revision of the workload estimates, new computer files were built to reflect the impact of each of the 26 automation options. Subsequently, the one- and two-crewmember models were exercised using the new files to predict workload associated with each of the 26 individual automation options and 16 different combinations of the individual automation options.

Table 3 presents results from exercising the one-crewmember workload model with the five individual automation options that produced the greatest reductions in workload. The Hover Hold and Automatic Sight Alignment options ranked highest with a 41.8% and 33.5% reduction in overload conditions and a 41.7% and 30.1% reduction in component overloads, respectively.

Table 4 presents results from exercising the two-crewmember model with the five automation options that produced the greatest reductions in workload. The Automatic Sight Alignment and Automatic Target Tracking options ranked highest with a 37.2% and 32.6% reduction in overload conditions and a 39.6% and 35.8% reduction in component overloads, respectively. The highest ranking option in the one-crewmember analysis, Hover Hold, reduced no overload conditions in the two-crewmember analysis.

Table 5 presents results from exercising the one- and two-crewmember models with a combination of all 26 individual automation options. The combination of 26 automation options reduced overload conditions 96.2% and component overloads 97% in the one-crewmember analysis. Reductions in psychomotor component overloads contributed the most (62%) to the reduction in total component overloads. The combination of 26 automation options reduced all of the overload conditions and component overloads in the two-crewmember analysis<sup>2</sup>. Reductions in visual and cognitive component overloads contributed more (39.6% and 32.1%, respectively) to the reduction in total component overloads than the reductions in psychomotor component overloads (28.3%).

Table 3. Workload Reduction From Five Highest Ranking Automation Options, One-Crewmember LHX Configuration

|                                                                   | % Reduction<br>in Overload<br>Conditions | % Reduction in<br>Total Component<br>Overloads |
|-------------------------------------------------------------------|------------------------------------------|------------------------------------------------|
| Automation Configuration                                          | N = 263                                  | N = 336                                        |
| Hover Hold                                                        | 41.8                                     | 41.7                                           |
| Automatic Sight Alignment                                         | 33.5                                     | 30.1                                           |
| Automatic Target Tracking                                         | 16.0                                     | 19.6                                           |
| Voice Recorder for Message Entry<br>During Low Workload Intervals | 5.7                                      | 4.8                                            |
| Automatic Updating of Position                                    | 5.3                                      | 5.7                                            |

Table 4. Workload Reduction From Five Highest Ranking Automation Options, Two-Crewmember LHX Configuration

|                                                                 | % Reduction<br>in Overload<br>Conditions | % Reduction in<br>Total Component<br>Overloads |
|-----------------------------------------------------------------|------------------------------------------|------------------------------------------------|
| Automation Configuration                                        | N = 43                                   | N = 53                                         |
| Automatic Sight Alignment                                       | 37.2                                     | 39.6                                           |
| Automatic Target Tracking                                       | 32.6                                     | 35.8                                           |
| Automatic Updating of Position                                  | 18.6                                     | 20.8                                           |
| Automatic Maneuver NOE                                          | 16.3                                     | 13.2                                           |
| Automatic Display of Location<br>Relative to Selected Waypoints | 11.6                                     | 15.1                                           |

<sup>2</sup>Results from another analysis indicated that all of the overload conditions in the two-crewmember baseline LHX model could be eliminated with a combination of only nine automation options (Aldrich, Szabo, & Craddock, 1986).

Table 5. Workload Reduction From a Combination of 26 Automation Options:  
One- and Two-Crewmember Analyses

| Automation Configuration                                               | % Reduction<br>in Overload<br>Conditions | % Reduction<br>in Component<br>Overloads | Relative Contribu-<br>tion to Overload<br>Reductions (%) |              |               |
|------------------------------------------------------------------------|------------------------------------------|------------------------------------------|----------------------------------------------------------|--------------|---------------|
|                                                                        |                                          |                                          | V                                                        | C            | P             |
| A Combination of<br>26 Automation Options--<br>One-Crewmember Analysis | N = 263<br>96.2                          | N = 336<br>97.0                          | N=79<br>23.3                                             | N=54<br>14.7 | N=203<br>62.0 |

|                                                                        |        |        |        |        |        |
|------------------------------------------------------------------------|--------|--------|--------|--------|--------|
| Automation Configuration                                               | N = 43 | N = 53 | N = 21 | N = 17 | N = 15 |
| A Combination of<br>26 Automation Options--<br>Two-Crewmember Analysis | 100.0  | 100.0  | 39.6   | 32.1   | 28.3   |

The results presented in Tables 2 through 5 demonstrate how the models can be used to conduct comparability analyses of operator workload for various crew and automation configurations. Similar analyses will be conducted for automation options being proposed for the AH-64A, UH-60A, and CH-47D aircraft.

#### RESEARCH REQUIRED TO VALIDATE THE WORKLOAD PREDICTION MODELS

Phase 4 consists of research required to validate the workload parameters used to develop the models and the workload predictions yielded by the models. Workload parameters that require validation include the:

- workload ratings assigned to each task,
- total workload estimates for concurrent tasks,
- estimated time required to perform each task,
- threshold for excessive workload,
- temporal relationships among tasks, and
- sequential relationships among tasks.

Specific predictions yielded by the models that require validation include the four indices of excessive workload described above.

A research plan (Aldrich & Szabo, 1986) describes the research required to validate the LHX workload prediction model. Although the research plan was developed specifically for the LHX, it can also guide research required to validate the AH-64A or UH-60A workload prediction models.

The validation research consists of three phases. During Phase 1, the reliability of the workload rating scales and the workload predictors are established. During Phase 2, validation data are collected through a series of studies employing part-mission and full-mission simulation. During Phase 3, the results from Phases 1 and 2 are used to refine the workload prediction model. Each of the three phases is described briefly below.

#### Establish the Reliability of the Workload Rating Scales

To accomplish this objective two surveys are required. The first survey presents pair comparisons of the verbal anchors for each workload rating scale to SMEs. The SMEs must choose the anchor in each pair that imposes more attentional demand. The survey results indicate the degree of agreement among the SMEs and also can be used to produce

equal-interval scales (Engen, 1971) to replace the ordinal scales that were used in the original workload analysis.

The first survey has been conducted for the LHX, AH-64A and UH-60A workload prediction models. In the case of the LHX and AH-64A, a consensus set of verbal anchors was developed for each of the five workload component scales. A survey instrument, comprising all pair comparison combinations of the consensus verbal anchors from each workload rating scale, was produced. The pair comparison survey was mailed to 71 human factors researchers and practitioners who are SMEs in workload research. The data from 38 completed surveys were used to develop each rater's rank order judgments of the verbal anchors. The rank ordered judgments were analyzed using Kendall's Coefficient of Concordance (Siegal, 1956) to assess the degree of agreement among the SMEs. The Coefficients of Concordance for the five scales are as follow:

- Visual - .39,
- Auditory - .46,
- Kinesthetic - .38,
- Cognitive - .69, and
- Psychomotor - .47

All of the above Coefficients of Concordance are significant at the .001 level, indicating a degree of consensus among the SMEs.

Bierbaum, Szabo, and Aldrich (1987) performed a similar analysis for the UH-60A workload component rating scales. They developed a pair comparison survey and personally presented the matched pairs of verbal anchors for the visual (both naked eye and night vision goggles), auditory, cognitive, and psychomotor workload component scales<sup>3</sup> to UH-60A IPs from the UH-60A AQC at the USAAVNC, Fort Rucker, Alabama. The data were used to develop each rater's rank order judgments of the verbal anchors. The Coefficients of Concordance for the five scales are as follow:

- visual, no goggles - .25 (19 IPs),
- visual, with night vision goggles - .18 (19 IPs),
- auditory, - .32 (14 IPs),
- cognitive, - .45 (11 IPs), and
- psychomotor, - .46 (15 IPs).

Although these Coefficients of Concordance are smaller than those computed from the LHX and AH-64A data, they also are significant at the .001 level. Thus, the coefficients indicate some degree of agreement among the IPs who provided the ratings.

The second survey has not yet been developed. It will ask SMEs to use the verbal anchors in the workload scales to rate the short descriptors of visual, auditory, kinesthetic, cognitive, and psychomotor workload components for each task in the model. Correlational techniques will be used to evaluate the interrater reliability of the workload ratings.

#### Employ Flight Simulation Research to Validate the Workload Prediction Model Parameters

Part-mission and full-mission simulation experiments will be required to validate the workload estimates produced by the models. For the part-mission simulation, mini-scenarios will be generated by selecting concurrent and sequential tasks from the mission and task

<sup>3</sup>The survey did not include verbal anchors from the kinesthetic scale because the analysts doubted that IPs would be able to distinguish between levels of attentional demand for the kinesthetic verbal anchors.

analysis. For the full-mission simulation, a composite mission scenario will be developed by selecting segments from the mission and task analysis.

The part-mission simulation will be conducted using a repeated measures experimental design in which each subject will fly the mini-scenarios multiple times. Results will be analyzed to assess the correlation between the workload model predictions and measures of the operators' performance on the concurrent and sequential tasks. The correlation coefficients will serve as the primary measure of how accurately the workload predictions forecast excessive workload at the task level of specificity. To assess the validity of the time estimates used in the model, the actual amount of time required to perform the various tasks in the mini-scenarios will be compared with the times estimated during the task analysis. The sequential relationships among the tasks will be evaluated by noting the subjects' ability to progress through the mini-scenarios following the sequence of tasks specified by the model.

During the full-mission simulation experiments, each trial will start at the beginning of a composite scenario and continue without interruption to the end. Analysis of results will include all of the analyses performed during the part-mission simulation data analyses. In addition, an analysis will be performed to assess the effects of inserting secondary tasks into the composite mission scenario.

The planned experiments have not been conducted because a flight simulation facility capable of supporting the part-mission and full-mission simulation studies has not been available. However, the new Crew Station Research and Development Facility (CSRDF), located at the Army's Aeroflightdynamics Directorate, NASA Ames, California recently procured a high-technology generic flight simulator that is ideally configured for validating the LHX workload prediction model. A high fidelity AH-64 flight simulator at McDonnell Douglas Helicopter Company or an Army AH-64A Combat Mission Simulator may become available for performing research required to validate the AH-64A workload prediction model.

#### Refine the Workload Prediction Models

Refinement of the workload prediction models has been on-going since the original LHX workload prediction models were completed. Improvements introduced during the development of the AH-64A model include:

- a model of the entire AH-64A combat mission, from preflight through postflight,
- a more granular mission/task/workload analysis at the switch and display element level of specificity,
- development of a scale for rating the kinesthetic workload component of mission tasks,
- expansion of the existing 4-point scale to a 7-point scale for rating the auditory component of mission tasks,
- categorization of discrete tasks into discrete fixed and discrete random tasks,
- analysis of visual workload component specifiers, internal viewing vs. external viewing, for identifying possible visual workload clashes for concurrent operator tasks,
- analysis of psychomotor workload component specifiers, left hand vs right hand, for identifying possible psychomotor workload clashes for concurrent operator tasks, and

- a listing of the type of switch for each task that involves a switch operation.

Improvements introduced during the development of the UH-60A workload prediction model include:

- development of a visual-aided workload component scale for rating visual workload while using night vision goggles, and
- development of equal-interval rating scales to replace the ordinal scales in the LHX and AH-64A workload models.

During the validation research, additional refinements will occur. The data from the pair comparison survey will be used to produce equal-interval rating scales to replace the ordinal scale values in the LHX and AH-64A data bases. The models will be exercised to produce refined workload predictions based upon the new scale values.

As the part-mission and full-mission simulation results are analyzed, additional refinements will be made to the workload prediction models. The researchers will make necessary corrections to the workload estimates, time estimates, and decision rules. Refined workload predictions will be produced using the empirically derived workload estimates and time values.

#### DISCUSSION AND CONCLUSIONS

The workload prediction methodology described above provides a systematic means for predicting human operator workload in advance of system design or system modifications. This section of the chapter (a) discusses some of the weaknesses and strengths of the methodology so that the reader may better judge the value of the workload prediction models, and (b) offers some conclusions for the reader to consider.

##### Methodological Weaknesses

In all of the workload analyses described in this chapter, the workload estimates assigned during the mission/task/workload analysis phase are the basic units of analysis. The greatest weakness in the methodology stems from the subjective nature of these estimates.

As previously described, the workload estimates consist of numerical values assigned to the sensory, cognitive and psychomotor components of each task. The assigned estimates are derived by comparing verbal descriptors of the tasks with verbal anchors judged to represent increasing levels of attentional demand. Until the scales are demonstrated to be both reliable and valid, any results from exercising the models can be questioned.

Another methodological weakness exists in the procedure that sums the subjective values of the task workload components to derive total workload estimates for a given component of concurrent tasks. In the LHX and AH-64A analyses, the subjective values are clearly ordinal. Summing ordinal values to derive total estimates is a questionable procedure. The development of interval scales will eliminate this weakness.

A related methodological weakness stems from the treatment of each of the different types of workload components, (i.e., visual, auditory, kinesthetic, cognitive, and psychomotor) as separate and independent entities. It seems doubtful that, in reality, psychomotor workload can exist independently of concurrent cognitive and visual workload.

The analysts' decision to designate a total value of "8" as the threshold for identifying sensory, cognitive, and psychomotor overloads

represents another subjective aspect of the research methodology. The selection of "8" is based solely upon the rationale that "7" is the upper limit of human capacity in the three workload modalities. Thus, it can be argued that the decision to use the value of "8" as the criterion for defining component overloads is an arbitrary one.

#### Methodological Strengths

The methodological weaknesses, considered by themselves, may lead the reader to question whether the methodology offers any advantages. However, certain strengths are believed to compensate for any impact that the weaknesses may have when applying the methodology to system design questions. The primary strength is that the methodology produces conservative estimates of workload.

First, whenever possible the decision rules are written to delay crew support functions on the timeline so that they will not conflict with high workload flight control and mission functions. In addition, the duration of flight control functions is extended so that all concurrent tasks can be presented on the timeline. To the extent that the function and task times are extended, the predictions of overload conditions and component overloads, produced by the stress of limited time are minimal.

Second, the criterion used to define excessive workload produces conservative estimates of component overloads. The methodology does not distinguish between varying degrees of overload when the sum of the ratings exceeds the threshold value of "8". The criterion value of "8" also precludes the recognition of instances in which a lower value may represent an overload condition. For example, a situation in which each of two or more workload components has a workload estimate of "6" may constitute a more critical overload condition than a situation in which only one component has a value of "8" or higher. In defining overload conditions, the methodology does not consider the total estimate for all three workload components.

A third way in which the predictions of excessive workload are conservative is that they predict workload under ideal operating conditions. The methodology does not consider increases in workload that will occur if mission performance is degraded due to visual obscuration, malfunctioning subsystems, or enemy activity. Obviously, such degradation would increase the workload beyond the level predicted by the present models.

A second major strength of the methodology is that it is designed to permit refinement during the analyses. Specifically, the methodology provides a means for refining the estimates of both workload and time as additional information becomes available. The workload estimates can be revised by assigning new verbal descriptors and numerical estimates to the workload components for each task; the timeline estimates can be revised by writing new decision rules.

#### Conclusions

The workload prediction methodology described above provides a systematic means of predicting human operator workload in advance of system design. The methodology predicts the attentional demand associated with the sensory, cognitive, and psychomotor components of individual and concurrent operator tasks. The workload predictions are computed and displayed on half-second timelines for both single- and dual-crew configurations. The workload estimates can be revised to predict the impact of (a) different crew configurations and (b) various automation options being considered during system design and system modifications.

In addition, the research methodology provides information for identifying emerging system personnel, manning, and training requirements. By assisting in the identification of these requirements, the methodology provides a means of developing early estimates of system personnel and training costs. The personnel and training cost estimates can then be factored into trade-off studies conducted during the early stages of system development.

#### REFERENCES

- Aldrich, T. B., Craddock, W., & McCracken, J. H. (1984). A computer analysis to predict crew workload during LHX scout-attack missions (Draft Technical Report ASI479-054-84[B], Vol. I, II, III). Fort Rucker, AL: Anacapa Sciences, Inc.
- Aldrich, T. B., & Szabo, S. M. (1986). Validation of the LHX one-crewmember workload prediction model (Draft Technical Memorandum ASI678-202-86[B]). Fort Rucker, AL: Anacapa Sciences, Inc.
- Aldrich, T. B., Szabo, S. M., & Craddock, W. (1986). A computer analysis of LHX automation options and their effect on predicted workload (Draft Technical Report ASI479-063-85[B]). Fort Rucker, AL: Anacapa Sciences, Inc.
- Bierbaum, C. R., Szabo, S. M., & Aldrich, T. B. (1987). A comprehensive task analysis of the UH-60 mission with crew workload estimates and preliminary decision rules for developing a UH-60 workload prediction model (Draft Technical Report ASI690-302-87[B], Vol. I, II, III, IV). Fort Rucker, AL: Anacapa Sciences, Inc.
- Engen, T. (1971). Psychophysics II: Scaling methods. In J. W. Kling and L. A. Riggs (Eds.), *Experimental psychology* (3rd ed.), pp. 51-54. New York: Holt, Rinehart, and Winston.
- McCracken, J. H., & Aldrich, T. B. (1984). Analyses of selected LHX mission functions: Implications for operator workload and system automation goals (Technical Note ASI479-024-84). Fort Rucker, AL: Army Research Institute Aviation Research and Development Activity.
- Siegal, A. I., Madden, E. G., & Pfeiffer, M. G. (1985). Task and training requirements for advanced Army helicopters (U.S. Army Research Institute Research Note 85-105). Fort Rucker, AL: Applied Psychological Services, Inc.
- Siegal, S. (1956). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.
- Szabo, S. M., & Bierbaum, C. R. (1986). A comprehensive task analysis of the AH-64 mission with crew workload estimates and preliminary decision rules for developing an AH-64 workload prediction model (Draft Technical Report ASI678-204-86[B], Vol. I, II, III, IV). Fort Rucker, AL: Anacapa Sciences, Inc.
- Taylor, R. R., & Poole, R. (1983). OH-58D MEP descriptions and workload analysis (Report No. 406-099-063). Fort Worth, TX: Bell Helicopter Division Textron.
- Wickens, C. D. (1984). *Engineering psychology and human performance*. Columbus, OH: Merrill.



# APPLICATIONS OF HUMAN PERFORMANCE MODELS TO SYSTEM DESIGN

Edited by

**GRANT R. McMILLAN**

*Harry G. Armstrong Aerospace Medical Research Laboratory  
Wright-Patterson Air Force Base, Ohio*

**DAVID BEEVIS**

*Defence and Civil Institute of Environmental Medicine  
Downsview, Ontario, Canada*

**EDUARDO SALAS**

*Naval Training Systems Center  
Orlando, Florida*

**MICHAEL H. STRUB**

*U.S. Army Research Institute for the Behavioral and Social Sciences  
Fort Bliss, Texas*

**ROBERT SUTTON**

*Royal Naval Engineering College  
Manadon, Plymouth, Devon, United Kingdom*

and

**LEO VAN BREDA**

*TNO Institute for Perception  
Soesterberg, The Netherlands*

(1988)



**PLENUM PRESS • NEW YORK AND LONDON**  
Published in cooperation with NATO Defense Research Group