The Measurement of Operator Workload in the Mark IVA DSCC Monitor and Control Subsystem

M. LeMay
Montclair State College, New Jersey

R. L. Chafin
Control Center Operations

E. E. Hird
Operations Sustaining Engineering Section

An operator workload measurement methodology is presented which will be used in support of the Mark IVA Operational Test and Evaluation (OT&E) Plan. Three operator workload measures are suggested: operator ratings, primary task work measures, and information processing time measures. A method of validating the workload measures using secondary task work measures is presented. We can expect that operations testing using these measures will assist in establishing the time required to perform essential operational activities and will indicate high risk operations areas due to potential operator overload.

I. Introduction

The goal of the Mark IVA DSN design is to provide a single cost-effective ground tracking and data acquisition network that will be capable of supporting both deep space and highly elliptical earth orbiter missions. To achieve this goal, existing tracking stations at Canberra, Goldstone, and Madrid will each be centralized so that the necessary antennas will be collocated near the signal processing equipment, which will be housed within a single building. Operating personnel will be reorganized and retrained to the extent necessary for them to be effective in their new roles.

Typically each of the present day Mark III DSN or GSTDN tracking stations is operated separately by its own staff. The Mark IVA DSN design will reduce the total amount of centralized equipment needed in conjunction with the antenna group to be developed at each 64-meter antenna site either by the relocation of an existing 34-meter antenna or by new construction.

A station supervisor in a signal processing center will be able to assign tasks to a control room crew consisting of a Complex Monitor and Control (CMC) operator working together with three Link Monitor and Control (LMC) operators. The centralized arrangement of the complex will permit a chosen level of tracking activities to be supported while the requirements for personnel and redundant equipment will have been reduced.

The CMC operator will be responsible for allocating assignable equipment resources to a link and for monitoring and controlling the global equipment that provides services to the complex. An LMC operator will be responsible for the acquisition of tracking data during a scheduled project support activity.

The LMC operator workload will be offset in the initial implementation by a combination of equipment automation and some additional roving operators who will perform specific tasks on manually operated equipment in the equipment room in response to LMC operator directions.

This article reports on a project whose purpose was to develop a method of measuring control room operator workload. This methodology is an essential part of the OT&E plan for Mark IVA operations testing. In an effort to estimate operator working times on the Mark IVA DSCC Monitor and Control Subsystem (DMC) a series of operations sequence diagrams (OSDs) have been drawn up. These are detailed task descriptions with estimates of the amount of time needed to complete each subtask by novice, experienced, and expert operators. These estimates are added to obtain a prediction of the amount of time needed for completion of a whole task such as a VLBI clock sync prepass. These times will be used as part of an effort to assess the actual workload on individual operators during task performance. This is necessary since it is likely that the occurrence of some events, particularly alarms, will result in a sudden, significant increase in workload which may, in turn, increase operator error rate, with an attendant risk of data loss. Minimizing the risk in the complex user-computer interface of the DMC is the purpose of the present effort.

Although much attention in recent literature has been given to the user-computer interface (Refs. 1, 2, and 3), most of it has not been aimed at gauging the effect of an increase in workload in an on-line, interactive system, with the possible exception of studies of pilot and aircrew performance, and this is a somewhat unique case of a system which is only partly, though increasingly, an information processing system. Moreover, most of the studies cited in an extensive search (Ref. 3) have been done in existing systems. The present problem is to assess workload on a system which is still in the design stage and to predict its effect on system performance. The measures developed to do this can be implemented in an operational test, which will show where human operator overload difficulties are likely to arise.

There is extensive literature on the measurement of workload (Refs. 3 and 4), and many possible measures have been proposed. These have been classified into four general categories (Ref. 3): operator ratings, spare mental capacity, primary task performance, and physiological measures.

A method of evaluating workload assessment measures based on this classification, also presented in Ref. 3, was applied to the operator's tasks as outlined in the OSDs, and a combination of operator ratings and primary task performance measures was chosen, with a spare mental capacity technique being used to appraise the validity of the measures. Physiological measures were not considered because they are expensive to implement, may be intrusive on the operator, and are not practical for ordinary on-line application. Some characteristics of the chosen measures are considered as follows:

A. Operator Ratings

Previous studies have shown this to be a reliable measure which is consistently related to performance (Refs. 5 and 6). It is relatively easy to obtain, and has the "face validity" of getting the operator's own estimate of the subjective phenomenon of mental workload. With careful attention to design (Refs. 7 and 8), rating scales yield quantitative results which may be used to gauge "crossover" points beyond which the chance for error may increase greatly or other system parameters may be affected.

B. Primary Task Performance

This is perhaps the most direct way to assess workload, since it involves the measurement of the effect of workload on the performance of the task of interest. The time needed to start and/or complete the task is the most relevant (Refs. 8 and 9), although a number of errors may also be considered. These measures must be carefully designed to reflect increasing workload, however, since human operators tend to adapt to the task and hold their performance constant over a broad range of conditions. Thus, operators' initial starting times for a particular subtask are generally taken as a measure of their strategy for dealing with increased load. The starting time reflects the speed at which incoming subtasks are processed mentally (Ref. 10), and is a linear function of the amount of information processed by the operator.

Additionally, successful completion of a particular subtask or set of tasks in a stated amount of time may itself be considered a measure of workload (Ref. 11), and putting increasingly stringent time requirements on the operator until task completion becomes impossible may reveal operator strategies as well as workload limits for particular subtasks. Therefore, both a measure of time to start (reaction time) and time to complete a subtask will be used to assess operator workload.

C. Secondary Task Measures

Since human operators tend to hold their task performance constant under increasing workload, primary task performance measures may change very little in the normal course of operation. If, however, a secondary task is introduced, so that the operator is required to divide attention between two tasks, then the secondary task may be considered a measure of "spare mental capacity" (Ref. 4) and, as the demands of the primary task increase, performance on the secondary task decreases, so it is an indirect measure of operator workload. This is the standard way of using a secondary task in workload measurement.

For on-line, operational measurement, the imposition of a secondary task is impractical, since it could interfere with normal operations. However, a variation of the secondary task technique can be used to validate the primary task measures in a trial or pilot study. Such a study is necessary in order to demonstrate that the workload measures used do, in fact, increase with increasing objective workload and that this in turn results in performance degradation and possible data loss.

For this purpose, the secondary task should be similar to tasks that would normally be performed by the operator, and it should be possible to make the task more or less time consuming. Log keeping in writing and by voice is such a task. It, in turn, will affect primary task performance as its requirements result in increased workload (Ref. 9). Then the primary task workload measures should rise, and this will result in a decrement in system performance.

Three specific ways of implementing rating and primary task workload measures are described below, along with a secondary task technique for validating the measures.

II. The Measures

Three measures are proposed for use in gauging operator workload in the DMC. They are interrelated, and should be collected simultaneously on a single task. They are:

(1) Sequential operator ratings of workload on subtasks.

(2) The ratio of the time required to the time available to do a particular subtask.

(3) Information processing time, or the time taken to initiate the physical (keystroke) portion of a subtask.

The latter two measures are keyed to the operator rating measure, which is taken at the end of each subtask and is recorded with a short free-form input (FFI).

A. Operator Ratings

1. Procedure. A single rating of workload, on a scale from 1 to 7 with a 7 meaning the highest workload, will be obtained from each operator during the performance of a task. The instructions to the operator (reproduced in Fig. 1) are given only once. Thereafter, at the completion of each subtask, the words "difficulty rating" and "time pressure rating," will appear in the SOE, and the operator will respond with an FFI text message and a rating of from 1 to 7.

A subtask is defined as a listed task with an estimated performance time in an operations sequence diagram (OSD). This will result in the operator's giving a rating at approximately 2-4 minute intervals, depending on operator experience.

The operator rating will be recorded on the operations log, along with the time at which it was made. This will provide a record of the time at which each subtask is finished (to be used in the other two measures described below), as well as of the operator's rating.

2. Analysis. The procedure will result in operator ratings on each of a number of subtasks, so that those points in the task at which overload occurs can be quickly located.

The operator ratings are obtained from the operations log and classified according to subtask and operator experience. Therefore, there should be some means of identifying the particular subtask (e.g., "Display and Compare APA Predicts against VLBI Source Table") on the log. The classification of operator experience (novice, intermediate, or expert) should be done by an experienced supervisor, with a set of guidelines taking into account the amount of time on the job and previous experience, if any.

Then the mean rating can be obtained for each subtask and each operator classification, and plotted as a function of task, resulting in a graph similar to that illustrated in Fig. 2. The peaks on such a graph will show the subtasks on which the operator is working at a higher subjective workload. In conjunction with the measures described below, this will indicate tasks on which errors might be expected to occur in the event of a sudden increase in workload such as might result from an event notice or alarm.

In this type of testing, the differences in ratings obtained for each subtask may be relatively small; the differences among the three types of operators may also be small but of interest. In order to test the significance of the differences, the rating data should be subjected to an analysis of variance (ANOVA), which shows whether or not there are real differences among tasks and/or operators, and whether there is an interaction between them, i.e., whether some tasks are perceived as having a higher workload by novice operators, and others by experienced operators, etc.

In addition, after the completion of a task, such as a Delta DOR Prepass operation or 64-meter antenna configuration, operators will rate several aspects of the job as a whole on a form like the one illustrated in Fig. 3. This is a very simple form, and can be analyzed easily, whether it is done on the log or by hand. Mean ratings on each of the first 10 questions should be obtained for each operator class. These should be examined individually as indicative of problem sources perceived by operators. The open-ended question 11 should also be used in the same way. The questionnaire will complement the subtask ratings by providing a general operator rating of workload in question 1. This rating should be correlated with the mean of the subtask ratings for each subject (i.e., x_i = mean of subtask ratings for S_i; Y_i = whole task workload rating from question 1 for S_i). It will also serve as an indication of problems perceived by operators in particular sub*systems*, as opposed to the problems of particular sub*tasks* pointed out in the repeated workload ratings obtained after each subtask.

B. Primary Task Workload Measure

1. **Procedure.** This measure will be similar to standard system reliability workload measures, in that it will consist of a ratio of time required to time available. Thus, workload is defined as:

$$W = T_r/T_a$$

where

W = workload

T_r = time required by the operator to perform a specific subtask

T_a = time estimated in the OSDs for that particular subtask.

Since we are estimating mental workload, there may be no obvious action recorded on the operations log to signal the start and end of a subtask. Therefore, the operator will be asked to signal the start of each task with a FFI which will be used to time that task on the log.

2. **Analysis.** The time between the FFI messages which indicate the start of a subtask and the operator's rating of that task can be taken as the actual or required time to perform the subtask. This time, T_r, will be available in printout form along with the operator ratings. It can then be compared with the times estimated in the OSDs for that subtask.

If subtasks are always performed in the proper sequence, particular tasks can be identified on this basis. However, an identifying code for each subtask, either a mnemonic or a numeral, would insure that the proper task was being considered with its time and rating, and would aid data analysis.

The resulting workload measure for each subtask will be analyzed in a fashion similar to that used for the ratings. Mean workload will be plotted as a function of task, with separate curves for each operator classification, and an analysis of variance performed on the data to determine if different subtasks yield significantly different workloads. In addition, the correlation between the workload measure and the operator rating for each subtask should be obtained.

If the obtained mean workload for a particular task is greater than 1, then the time needed for the task is greater than that originally allowed in the estimate. If it is above 0.7 or 0.8, it is approaching the operator's maximum capacity. In either case, if the estimates are to be considered as the time allowed for the task, then some provision must be made to relieve operator workload on these subtasks, perhaps by task rearrangement, extra help from personnel or computer, or simply increasing the time allowed. Otherwise, there will be an increased risk of operator error occasioned by sudden increases in workload, such as the occurrence of an event or alarm requiring operator action during an already overloaded subtask.

C. Information Processing Time

1. **Procedure.** The third value to be used does not measure workload directly, but is an attempt to measure the time required for mental processing of the information needed to perform a control function.

Many of the subtasks in the OSDs require only *monitoring* functions; in this case, since there is no overt action, but only information processing (e.g., comparing source time and ID against VLBI source table for compatibility), this information processing or "think" time will be the same as T_r (time required) in the second measure. However, on most subtasks, an overt response is required. The response may be part of the normal SOE or be contingent on the detection of some anomaly such as a high system temperature value. However, the majority of tasks do require some sort of physical action, and the time to start such action is called a choice reaction time (CRT), since the operator must choose the action to be taken from among a number of possible actions. It varies directly with the number of choices or actions, and is a measure of uncertainty, or the amount of information the operator must process. The longer the CRT, the more complex the information processing task and the greater the workload. Thus, information processing time, as measured by CRT, is also a measure of workload.

One task follows another immediately in the OSDs. However, this may not always be the case in practice, since there may be interruptions or operator recovery time from the previous task. It is reasonable to include this extra time in the direct workload measure (the second measure), since it becomes part of the workload and adds to it. However, it is obviously not part of the information processing time for a particular subtask. Therefore, it will be necessary, as for the second measure, for the operator to signal the start of each subtask with an FFI text message which can be used, in conjunction with the next action taken (keystroke or voice communication), in determining the information processing time.

2. Analysis. The time between the FFI text message which signals the start of a task and the next action is obtained from the operations log for each subtask for each operator. The mean of these values is obtained for each class of operators. This data is then subjected to the same type of analysis as before; i.e., mean information processing time is plotted as a function of task for each class of operators, so that those tasks which carry a heavy information processing load can be identified. The data is also subjected to a 3 (operator class) $\times$ n (number of tasks) $\times$ S (subjects) ANOVA, to determine the significance of the observed differences.

In addition, when all three measures have been obtained, a grand mean across all three classes of operators should be calculated for each measure on each task. If all three tasks are measuring workload, then there should be some degree of correlation among them. In fact, since the information processing time is a fraction of the required time in the workload measure, these two will probably be highly correlated. This makes the calculation of a multiple correlation coefficient inappropriate. However, it is appropriate to calculate the correlation coefficients between the operator rating and each of the other measures. These should be relatively high (above 0.5), which would indicate that all three measures are measuring different aspects of the same thing (workload) and would be evidence for the validity of the measures.

III. Validation of the Measures

The three proposed measures seem logical, and they are based on measures that have been used for similar purposes in the literature (Refs. 3 and 4). Nevertheless, their validity and reliability need to be evaluated for the use in the operational testing. It is possible, for example, that one or two of the measures will show no variation, even when operator workload is obviously markedly increased, whereas a minor change in the data collection method would yield usable results. This is a common occurrence in experimental trials with human operators. The only way to determine whether or not a measure of human performance will reflect variations in external working conditions is to run a "pilot study"; i.e., try the measure out using a small sample of people performing a relevant task.

In addition to demonstrating that the measure varies with task parameters, pilot studies can also be used to assess the effect of increasing workload on performance. The purpose of measuring operator workload, after all, is first to detect tasks on which the workload is so high that the task becomes impossible (this would be relatively rare) and second to detect those tasks or parts of tasks in which the workload is high enough so that, if a further burden is added in the form of, say, an unforeseen event, operator errors are likely to occur. If the measures are valid, i.e., if workload is really being measured, then they should increase with increasing task demand, and this should result in increased operator errors. The pilot studies, or trials, are designed to assess this.

The trials must be run before full-scale operational testing, both to allow for changes in the measures and because they will involve "loading" the operator with extra, or secondary, tasks until the error rate starts to increase, a technique that would be undesirable during full-scale operations testing.

1. Procedure. From three to six people, preferably those with some experience on the Mark III or a similar system, should serve as subjects for the trials. They need not be rated as novice, intermediate, or expert, since the purpose of the trials is simply to see whether the measures respond to increasing objective workload.

A task of about 30-45 minutes duration will be designed for the trials so as to be as similar as possible to the Mark IVA LMC operator's task, i.e., an interactive computer task. Although the LMC operator's tasks may be intermittent, with periods of high workload alternating with relatively light workload, the trial task should occupy the operator continuously.

The operators will perform the task under each one of three conditions. In the first condition, the primary task alone will be performed by the operator (low workload); in the second condition, a moderate amount of a secondary task, to be done at the same time as the primary task, will be assigned (medium workload); in the third condition, a still heavier amount of the secondary task will be assigned the operator (high workload). The purpose of the secondary task is twofold. First, it should demonstrate that as the secondary task increases the workload on the operator, the three workload measures rise, and second, performance on the secondary task should be negatively correlated with the workload measures, since, as primary workload goes up, the operator has less "spare mental capacity" to work on the secondary task. Both these effects serve as checks on the validity of the workload measures.

There will, of course, be no secondary task in the first condition. In the second condition, the secondary task will consist in the operators' keeping a written log of all of the operations and parameters. This should be done on a prepared form, similar to forms currently in use. In the third condition, in addition to keeping the written log, the operator will keep a voice log by reporting the data over the voice circuit to a receiving operator, similar to the NOCC operator. These two tasks have three characteristics: (1) they are designed to overload the operator in two increasing steps so as to produce an increase in the workload measures and possibly an increase in procedural lapses (errors of a type that would not result in data loss) on the part of the operator; (2) they are designed to be similar to secondary tasks normally performed by operators, so they will not be perceived as simply "busy work" that can be ignored; and (3) the amount of work involved in them can be manipulated in terms of the amount of detail required. Thus, if on first presentation the overloading tasks are found to be easily handled by the operator, and make no difference in the workload measures or in lapses, the amount of record keeping required should be increased until an effect is shown.

The operators should practice on the simulated system before the trials until their performance is asymptotic in terms of the amount of time they require to do the job. Then the trials with the three workload conditions can begin. Each person who serves as a subject should experience all three conditions, so that performance differences among conditions can be attributed to the effects of the workload variaton and not to individual differences. Each individual should experience the workload conditions in a different order chosen randomly from among the six possible orders of three conditions.

2. Analysis. The three workload measures described above will be taken for each individual for each subtask under each condition. Means will be calculated for each condition, and an analysis of variance run separately for each measure. This should demonstrate that the means of the workload measures increase significantly with the increase in imposed workload over the three conditions, thus providing evidence for the validity of the measures.

It would also be of interest to gauge the effect of workload on operator performance. Generally, operators compensate for increased workload by increased efficiency, so that errors which actually result in lost data are rare, and probably would not occur at all in a 30-45 minute trial. However, anomalies such as an incorrect input, failure to detect a significant reading on a display, or failure to understand a verbal communication are fairly common. They are usually corrected in a short time (so human operators are sometimes called "self-correcting") but they may significantly add to the workload itself and, more importantly, if the workload is already high, they may lead to further errors which may actually result in missed data. Therefore, an increase in operator anomalies with increasing workload must be taken seriously, even though the anomalies are corrected.

In the trials described here, then, all operator anomalies should be recorded in the course of data collection. They can then be subjected to the same type of analysis as the workload measures, i.e., an anlysis of variance which should demonstrate that anomalies increase as imposed workload increases. Further, a detailed examination of the subtask in which anomalies occur, and their workload measures, should yield useful data on the possible effect of workload on the probability of operator anomalies, on which some predictions may be made about the probability of data loss given a particular level of workload.

Some operator anomalies which should be considered are:

(1) Failure to notice a signal change or update

(2) Hitting the wrong input key

(3) Failure to find the source of an error message

(4) Voice communication failure (incorrect reception or transmission), including requests to "say again"

(5) Incorrect information recorded on the log

(6) Incorrect information transmitted to the voice log.

Furthermore, since such anomalies are inherently unpredictable, it is not possible to define them completely in advance, so that any anomaly that occurs during data collection should be recorded. This will involve some on-the-spot judgment on the part of the test conductor. The slightest anomaly should be recorded so as to increase the data pool.

If this procedure is followed, a useful and meaningful (in terms of probability of data loss due to operator error) measure or measures of human operator workload can be obtained. The measures can then be used in an operations test as a part of systems evaluation and periodically thereafter whenever system testing is needed.

IV. Implementation

The operator workload measurement methodology would use specially developed SOEs for operations testing. The SOEs would indicate when an operator rating is required. The rating will be input to the system by means of a text message which is timed tagged and recorded on the operations log. The analysis will use the data recorded on the operations log.

Validating the operator workload measurement with the secondary task method would be accomplished before the formal Mark IVA operational workload testing is begun.

V. Summary

Operator overload is identified as a potential problem in Mark IVA system operations. Three methods for measuring operator overload have been presented. The methods are: operator ratings, primary task work measures, and information processing time measures. A method of validating these three workload measures by secondary task performance measures is suggested.

These operator workload measures would be an integral part of the OT&E plan for operations testing of the DMC. Operations testing in this form would provide indications of where in the operations of the Mark IVA system we can expect system failures due to operator overload.

The operator workload measurement methodology described here allows the problem of operator overload to be addressed directly. It provides a general capability to measure operator workload in future systems.

References

1. Benbosat, I., Dexter, A., and Masubi, P., "An experimental study of the human/computer interface," *Communications of the ACM*, 1981, 24, 11, pp. 752-762 Nov. 1981.
2. Schneiderman, B., *Software Psychology: Human Factors in Computer and Information Systems*, Winthrop Publishers, Inc., Cambridge, Mass., 1980.
3. Wierwill, W. W., Williges, R. C., and Schiflett, S. G., "Aircrew workload assessment techniques," in Hartman, B. O., and McKenzie, R. E., *Survey of Methods to Assess Workload*. AGARDograph No. 246, Aug. 1979.
4. Moray, N., *Mental Workload: Its Theory and Measurement*. Plenum Press, New York, 1979.
5. McConnell, J. P., "An application of measurement methods to improve the quantitative nature of pilot ratings scales," *IEEE Trans. Man-Machine Systems*, MMS-10, pp. 81-92, 1969.
6. Hicks, T., and Wierwille, W., "Comparison of five mental workload assessment procedures in a moving-base simulator," *Human Factors*, 21, pp. 129-144, 1979.
7. Dyer, W., Matthews, J. J., Stalac, J. F., Wright, C. E., and Yudowitch, K., *Questionnaire Construction Manual Annex*, U. S. Army Research Institute for the Behavioral and Social Sciences, AD43012, July 1976.
8. Adams, J. A., "Issues in human reliability." *Human Factors*, 24, pp. 1-10, 1982.
9. Sternberg, S., "Memory-scanning: mental process revealed by reaction-time experiments," *American Scientist*, 57, 4, pp. 421-457, 1969.
10. Chiles, W. D., and Alluisi, E. A., "On the specification of operator or occupational workload with performance-measurement methods," *Human Factors*, 21, pp. 515-528, 1979.
11. Albanese, R. A., Mathematical analysis and computer simulation in military mission workload assessment," *Proceedings of the AGARD Conference on Methods to Assess Workload*, AGARD-CPP-216, AB-1-A13-6, April 1977.

We are interested in obtaining information about specific points in the sequence of operations where the workload or amount of effort required by you, the operator, approaches the level where the time allowed is not sufficient to get the job done. In order to do this, we will ask you for a series of ratings of the difficulty of the task and the amount of workload you experience on each of many of the small subtasks of a particular job. At frequent intervals while you are doing a job, a request for two ratings will appear in the SOE. At these times, please respond as fast as you can with the appropriate FFI text message which includes a value of from 1 through 7 to indicate your ratings of the difficulty of the task and the workload you have experienced since the last rating (or since starting the job). Use the following scale.

The task was:

(1) Extremely easy to do.

(2) Very easy to do.

(3) Easy to do.

(4) Neither hard nor easy to do.

(5) Hard to do.

(6) Very hard to do.

(7) Extremely hard to do.

1 2 3 4 5 6 7

increasing difficulty –

Difficulty rating

There was:

(1) Plenty of time to spare.

(2) Some time to spare.

(3) A little time to spare.

(4) Time allowed is just enough to get the job done.

(5) Hard to finish in the time allowed.

(6) Very hard to finish in the time allowed.

(7) Impossible to finish in the time allowed.

1 2 3 4 5 6 7

increasing time pressure –

Time pressure rating

Fig. 1. Instructions and scale for operator workload ratings

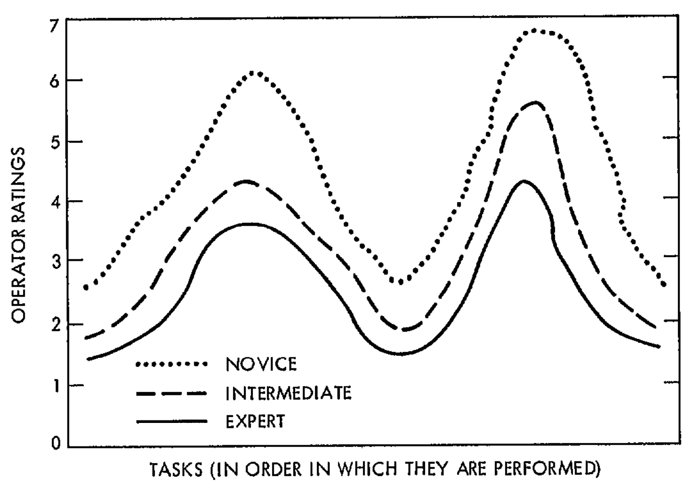

Fig. 2. Mean workload ratings for each subtask

The following questionnaire should be given once after the completion of a trial and/or after a reasonable familiarization period during operational testing.

Please indicate your rating of the following statements by writing a number from 1 to 7 in the space provided:

1 – Very strongly disagree.

2 – Strongly disagree.

3 – Disagree.

4 – Neither agree nor disagree.

5 – Agree.

6 – Strongly agree.

7 – Very strongly agree.

1. The workload required of the operators of this system is too high. ___
2. The system provides all the displays that are needed for the job. ___
3. In general, each display is adequate for its job. ___
4. The working group was able to respond to the demands placed on it. ___
5. Information can move freely and accurately among the LMC, roving, and NOCC operators. ___
6. It is easy to find the information needed in the system documentation. ___
7. System event reporting makes clear what action is to be taken. ___
8. Error messages describe each fault in sufficient explanatory detail. ___
9. Error messages make clear what action is to be taken. ___
10. There are too many meaningless messages on the screen. ___
11. There were other problems in operating the system (specify). ___

Fig. 3. Operator workload questionnaire