The Applicability and Suitability of Subjective Workload Measures for Implementation During Fixed Wing Military Flight

by A.C.Markou, B.Sc., C.D.Braby, B.SC., M.Sc., and H.C.Muir, M.A., Ph.d.

Applied Psychology Unit
College of Aeronautics
Cranfield Institute of Technology
Cranfield, Bedford MK43 0AL. England
College of Aeronautics Report No.9104
February 1990

The Applicability and Suitability of Subjective Workload Measures for Implementation During Fixed Wing Military Flight

by A.C. Markou, B.Sc., C.D. Braby, B.Sc., M.Sc., and H.C. Muir, M.A., Ph.d.

Applied Psychology Unit
College of Aeronautics
Cranfield Institute of Technology
Cranfield, Bedford MK43 0AL. England

ISBN 1 871564 13 1

£8.00

"The views expressed herein are those of the authors alone and do not necessarily represent those of the Institute"
This report summarises research conducted for the award of an M.Sc. degree to the first author.

Any enquiries concerning the content of this report can be made to the address below:

Androulla Markou,
Applied Psychology Unit,
College of Aeronautics,
Cranfield Institute of Technology,
Cranfield,
Bedford,
MK43 OAL.

Tel: 0234 750111 (ext. 2228/2229)
The research reported herein represents an initial investigation into the applicability and suitability of four commonly-used subjective measures of workload: the NASA-Task Load Index (NASA-TLX), Subjective Workload Assessment Technique (SWAT), Modified Cooper-Harper (MCH) and Bedford scales.

The study was conducted in two phases. The first phase assessed the applicability of the subjective measures to the measurement of workload in an air-to-air and air-to-ground military flight scenario. Military pilots from RAF bases participated in this phase, during which applicability indices for each measure were gathered with the use of a self-completion questionnaire. The second phase of the study assessed the suitability of the measures for implementation during fixed wing, military flight. Pilots with military experience performed an air-to-ground scenario on a "Genesis" simulator, during which workload ratings were gathered using all four workload measures. The primary purpose of the simulation exercise was to give the participants an opportunity to appreciate what in-flight implementation of the measures would involve, thus allowing suitability ratings to be given during a post-task interview.

The findings suggested that: (i) the unidimensional scales were more suitable for implementation during military flight than the multidimensional measures, (ii) of the four measures assessed, the MCH scale fared best on the criteria of both applicability and suitability, and (iii) the applicability and suitability of all the measures was highly dependent on the task situation (i.e. the particular manoeuvre, flight segment or flight scenario being performed).
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Method</td>
<td>4</td>
</tr>
<tr>
<td>Results</td>
<td>8</td>
</tr>
<tr>
<td>Discussion</td>
<td>13</td>
</tr>
<tr>
<td>Conclusions</td>
<td>15</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>16</td>
</tr>
<tr>
<td>References</td>
<td>17</td>
</tr>
<tr>
<td>Appendix A: The Bedford and MCH scales and dimensions of the SWAT and NASA-TLX</td>
<td></td>
</tr>
<tr>
<td>Appendix B: Applicability questionnaire, flight scenarios and covering letters</td>
<td></td>
</tr>
<tr>
<td>Appendix C: Counterbalanced design for presentation of workload measures</td>
<td></td>
</tr>
<tr>
<td>Appendix D: Revised SWAT and NASA-TLX dimensions used in current study</td>
<td></td>
</tr>
</tbody>
</table>
INTRODUCTION

DEFINITION AND MEASUREMENT OF WORKLOAD

The past decade has witnessed a marked increase in the interest of human factors practitioners in operator workload, an interest reflected in the growing number of workload studies documented in the literature. Nowhere has this increase been more apparent than in the aviation industry. The rapid advance of technology means that aircraft systems with immense capabilities can now be developed; the human operator of such systems can be presented with an almost infinite array of information. However, there is a limit to the amount of information the human operator can process at any given time. This presents what has come to be recognised as the primary limiting factor in the system (Detro, 1985; Eggemeier, 1980). The potential for exceeding the operator's processing and response capabilities has increased and poses a problem of serious consequences. Therefore the need for appropriate methods of assessing the level of workload to which the operator is subjected has assumed increased importance in recent years.

Workload is obviously an abstract concept; it is not an entity which can be handled and objectively examined. Consequently, no single, universally accepted definition of workload exists. However, Eggemeier (op cit.) has pointed out that most definitions do have a number of common elements, including "some characterization of the array of demands placed on the operator by the system, and some expression of the capacity or capability of the operator to deal with such demands as indicated by some aspect of performance" (page 669).

In much the same way that intelligence is regarded by psychologists as consisting of various sub-scales or dimensions (verbal, arithmetic etc.), so workload is also viewed by the majority of human factors practitioners as being "multidimensional" (Eggemeier, op cit.; Hart & Staveland, 1987; Reid, Shingledecker & Eggemeier, 1981). Unfortunately, the number and type of dimensions necessary for a complete description of workload is still a matter for debate (Eggemeier, op cit.; Vidulich & Tsang, 1986).

The difficulty in identifying the dimensions which contribute to workload is compounded by the fact that different tasks appear to "tap" different dimensions. This issue is associated with multiple-resource (or multiple capacity) theory, which states that instead of a single central "pool" of resources, humans possess several such pools, each with their own capacities and resource properties. Tasks which draw upon the same pool of resources will reveal performance decrements when carried out simultaneously (Wickens, 1984). Conversely, tasks which draw upon different pools will reveal no such decrements. It has been suggested that these separate resources can serve as the sought-after dimensions of operator workload (Derrick, 1988).
As the need for the study of workload has escalated, so the number of workload assessment techniques available has increased. These techniques are generally classified into one of several broad categories such as "performance", "physiological" or "subjective". An example of a performance measure is the secondary task measure, which requires an operator to perform a specified primary task and to use any spare capacity to simultaneously perform a secondary task. The decrement in performance of the secondary task is operationally defined as a measure of workload. One of the most common physiological measures is that of heart rate, which is generally found to increase under conditions of high workload (Roscoe, 1984). Subjective measures involve asking operators to provide an indication of how hard they believe they have been working. These measures have proved highly popular in the measurement of workload for various reasons: they are relatively inexpensive, convenient, easy to analyse and have high operator acceptance (Eggleston, 1984).

This study assessed the applicability and suitability of some of the most commonly-used subjective measures of workload:

- The unidimensional Modified Cooper-Harper (MCH) scale, introduced by Wierwille and Casali in 1983.
- The unidimensional Bedford Workload scale, developed by Roscoe and Ellis in 1984 (Roscoe, 1986).
- The multidimensional Subjective Workload Assessment Technique (SWAT), developed by Reid, Shingledecker and Eggemeier in 1981.
- The multidimensional NASA Task Load Index (NASA-TLX), developed at the NASA Ames Research Center (Hart & Staveland, *op cit.*)

**THE APPLICABILITY OF SUBJECTIVE MEASURES OF WORKLOAD**

The identification and measurement of the relevant human capabilities being used to perform a task has been referred to by some human factors practitioners as the criterion of diagnosticity (Gawron, Schfft and Miller, 1989). However, in the current study, the definition of this criterion was extended in order to include the identification and measurement of factors, such as stress and frustration, which may have an impact on the overall perception of workload in a particular task situation. The term "applicability", rather than diagnosticity, was therefore adopted.

A subjective workload measure is usually based on several fundamental factors, each of which may be relevant to differing extents across task conditions. As van de Graaff (1987) points out: "It is assumed that workload generally encompasses several components...It is not clear which components play a part in a specific situation, nor what the impact of each is upon the overall perception of workload" (page 9). Moreover, in the field of aviation workload, the majority of studies have been concerned with civil rather than military aviation.

* Appendix A includes a copy of each of the unidimensional scales and the dimensions on which the multidimensional measures are based.
Unfortunately, findings gleaned from workload studies in the civil aviation environment cannot readily be generalized to the military environment. Not only are military aircraft systems considerably more complex than their civil counterparts, the sorties and missions which military pilots are required to fly are many and varied. Scenarios typically flown in military aircraft include air-to-air and air-to-ground missions. Although both are offensive missions, the air-to-air scenario involves the attack of another aircraft during flight using missiles, whereas the air-to-ground scenario involves the attack of a ground target using bombs. Clearly the two scenarios may differ in the physical and mental human capabilities which they require. Thus, a workload measure which successfully assesses factors relevant, or applicable, to the workload imposed during one particular flight scenario may not do so during another.

The aim of the first phase of this study was therefore to assess the applicability of the NASA-TLX, SWAT, MCH and Bedford scales to the measurement of workload in fixed wing military flight.

The suitability of subjective measures of workload

The term "suitability" refers to the practical constraints associated with the implementation of a particular measure. When used for the assessment of pilot workload, subjective workload measures can be implemented during flight (in-flight implementation) or following a flight (post-flight implementation). Since some workload measures do not easily lend themselves to in-flight implementation, they are often administered after the flight. Unfortunately, findings reported by many researchers suggest that post-flight administration of a subjective measure may yield different ratings to in-flight implementation of the same measure (Corwin, 1989; Lidderdale, 1986).

Moreover, even if a measure is suitable for implementation during flight, the flight segment during which the measure can be implemented needs to be considered. In civil aviation, a routine flight can be divided into seven segments: "takeoff", "climb", "top of climb", "cruise", "top of descent", "approach" and "landing". Each of these segments provides a distinct "measurement window" for the assessment of workload. That is, a workload rating can generally be obtained for each of these segments. Such delineation of measurement windows is considerably more difficult to achieve in military aviation, where each scenario will differ in the number and type of segments. A workload measure which is suitable for use in civil aviation may therefore not be suitable for use in military aircraft. By the same token, a measure which is deemed suitable for implementation during one segment of a flight may not be suitable for use during another.

Thus the aim of the second phase of this study was to assess the suitability of the NASA-TLX, SWAT, MCH and Bedford scales for implementation during fixed wing military flight.
METHOD

PHASE I: THE APPLICABILITY OF SUBJECTIVE MEASURES OF WORKLOAD

Overview

A self-completion questionnaire was developed to assess the applicability of each subjective measure with regard to various segments of an air-to-air and an air-to-ground scenario; a task analysis of these scenarios was attached to the questionnaire.

Participants

Four test-pilots at the International Test Pilots School (ITPS), Cranfield, and ten tornado pilots from three RAF bases in Britain and two in Germany participated in the first phase of the study.

Applicability questionnaire

A self-completion questionnaire was developed to evaluate the applicability of the four workload measures*. Each of the measures under consideration was examined and the component dimensions were extracted.

The NASA-TLX ("Assessment Technique 1") is based on six dimensions; mental demand, physical demand, temporal demand, performance, effort and frustration. Each of these dimensions is defined in the manual. Thus the original definitions were used in the questionnaire, with only slight modifications in wording to aid concision.

The SWAT ("Assessment Technique 2") includes standard definitions for each of the three levels of three dimensions; mental effort load, time load and psychological stress load. Key words were combined into a single definition for each of the three dimensions which was included in the questionnaire.

In the case of the MCH ("Assessment Technique 3") and the Bedford ("Assessment Technique 4") scales, the process of isolating the basic factors was slightly more subjective. Although each is a unidimensional workload measure, and hence views a single rating of workload to be sufficient, examination of the structure and wording of the measures reveals several possible fundamental factors on which they may be based. Eventually, three factors were included in the questionnaire for the MCH scale; mental effort, task difficulty and nature of errors. Two factors were included for the Bedford scale; spare capacity and possibility of task completion.

Respondents were required to assess the applicability of each factor with respect to each of two flight scenarios which are typically carried out in military fixed wing aircraft; air-to-air (counter-air) and air-to-ground (close air support)*.

* The applicability questionnaire and flight scenarios are included in Appendix B.
A detailed task analysis of both these scenarios was reported by Hollister (1986) and was incorporated into the study. The complete flight scenarios were not employed as they were too lengthy. In addition, the "take-off and climb" and the "approach and landing" segments were virtually identical for the two scenarios; inclusion of both segments for each scenario would therefore have resulted in unnecessary repetition.

Pre-testing of the questionnaire was carried out on four test pilots at the ITPS.

Distribution of the questionnaire

Fifty-five copies of the questionnaire were posted to RAF bases in England and abroad. A covering letter was included for the Commanding Officer of each base, as well as for each respondent (see Appendix B).

PHASE II: THE SUITABILITY OF SUBJECTIVE MEASURES OF WORKLOAD

Overview

Participants were required to perform an air-to-ground scenario on a "Genesis" simulator, during which workload ratings were gathered using all four workload measures. The primary purpose of the simulation exercise was to give the participants an opportunity to appreciate what in-flight implementation of the measures would involve, thus allowing suitability ratings to be given during a post-task interview.

Participants

Sixteen male, fixed wing pilots participated in the second phase of the study. All pilots had military experience varying from 2 to 38 years, with an average of 13.5 years (sd=8.0).

Procedure

1) Initially, participants were informed of the aim of the study and familiarised with the workload measures; the researcher introduced each measure and explained how ratings are derived.

2) The "scale development" stage of the SWAT was performed. Participants were presented with a deck of 27 index cards, each card listing three workload descriptors, one of each workload dimension. Participants were instructed to place the cards in order, from lowest to highest workload.

3) Each participant then performed a "training" scenario on the simulator. This involved taking-off, climbing to 10,000 feet and orbiting once before landing.
4) Participants were briefed on the air-to-ground scenario they would be required to perform. This involved climbing to 10,000 feet, levelling off and entering a series of three orbits. After three orbits, the attack phase was negotiated, immediately after which the participants were required to climb back to 10,000 feet and enter another series of orbits. During the simulation, ground radar control instructions were delivered by the researcher.

5) The flight scenario was then performed on the simulator.

6) Ratings for each of the workload measures were taken during the second series of orbits, immediately following the attack phase. Participants were instructed to rate their workload during the attack phase, between the two orbit points. The order of presentation of the workload measures was counterbalanced (see Appendix C).

In the case of the two unidimensional scales (the MCH and the Bedford scales), the researcher read through the decision trees, addressing the questions to the participants. However, in order to avoid confusion and the need for repetition, the SWAT and NASA-TLX definitions were abbreviated and modified into questions which were addressed to the participants (see Appendix D).

During the SWAT rating process, the numerical ratings (1, 2 or 3) corresponding to the participants' responses were recorded by the researcher. These ratings corresponded to one of the combinations rated during the earlier "scale development" stage. Thus, the overall subjective workload score assigned to the task was obtained by locating the scale value previously assigned to that particular dimension combination.

During the NASA-TLX rating procedure, respondents are normally asked to place a tick on a scale for each dimension. However, due to various practical constraints (e.g. simulator controlled by right-hand stick) participants were instead asked to give a verbal rating for each dimension.

7) When the scenario was completed, the NASA-TLX weighting ("magnitude of load") procedure was conducted. Participants were presented with a deck of 15 randomly shuffled cards, each card listing a particular pairwise pairing of two of the six dimensions, and were asked to indicate which of the two dimensions on each card contributed more to the workload imposed by the task. The number of times each dimension was selected (the "weight") was noted and was multiplied by the corresponding rating to give a "weighted rating" for each dimension. The overall workload score was then calculated by dividing the sum of the weighted ratings by fifteen.

8) A semi-structured post-task interview was conducted by the researcher. Participants were asked to consider the suitability of each workload measure for implementation during an orbit in a fast jet. The level of workload associated with an orbit is normally fairly low. It was therefore assumed that, if a particular measure
was considered to be totally unsuitable for implementation during an orbit, it would probably not be suitable for implementation at any other point during the mission. The administration procedure to be considered was that of verbal presentation of the measures through the pilot's headset. Each participant was asked to assign a suitability index, of between zero (totally unsuitable) and 10 (highly suitable), to each measure.

The participants' opinions on post-task implementation of assessment measures and the use of a keypad for indicating workload levels during flight were also sought.

9) Participants were paid £10 per hour for taking part in the study.
RESULTS

PHASE I: THE APPLICABILITY OF THE SUBJECTIVE MEASURES

Completed questionnaires were coded as follows: Responses of "very applicable" received a score of 3, responses of "fairly applicable" received a score of 2, responses of "not applicable" received a score of 1.

Initially, the overall applicability index assigned by each subject to each measure was calculated; this index represented the average of all applicability ratings assigned to a particular measure across both segments of both scenarios.

The mean values and standard deviations of the overall applicability indices assigned to each measure are illustrated in table 1 below.

<table>
<thead>
<tr>
<th>Workload measure</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bedford</td>
<td>2.163</td>
<td>0.400</td>
</tr>
<tr>
<td>MCH</td>
<td>2.342</td>
<td>0.285</td>
</tr>
<tr>
<td>SWAT</td>
<td>2.092</td>
<td>0.213</td>
</tr>
<tr>
<td>NASA-TLX</td>
<td>2.179</td>
<td>0.217</td>
</tr>
</tbody>
</table>

TABLE 1: Means and standard deviations of the overall applicability indices assigned to each workload measure (N = 10)

The MCH scale received the highest average applicability index (2.342), followed by the NASA-TLX (2.179), the Bedford scale (2.163) and the SWAT (2.092). Statistical analysis (repeated measures ANOVA, multiple dependent t-tests and the Bonferroni inequality) revealed the MCH scale to be significantly more applicable than the NASA-TLX (p<.003) and the SWAT (p<.002). The Bedford scale was not significantly different to any other measure; this was probably due to the comparatively large spread of scores on this measure (std.dev.=0.4).

The average applicability index assigned to each measure was also calculated for each scenario. The means and standard deviations of these applicability indices for each scenario are shown in table 2 overleaf.
AIR-TO-AIR SCENARIO

<table>
<thead>
<tr>
<th>Workload measure</th>
<th>Mean</th>
<th>Std dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bedford</td>
<td>2.050</td>
<td>.422</td>
</tr>
<tr>
<td>MCH</td>
<td>2.217</td>
<td>.360</td>
</tr>
<tr>
<td>SWAT</td>
<td>2.000</td>
<td>.222</td>
</tr>
<tr>
<td>NASA-TLX</td>
<td>2.158</td>
<td>.259</td>
</tr>
</tbody>
</table>

AIR-TO-GROUND SCENARIO

<table>
<thead>
<tr>
<th>Workload measure</th>
<th>Mean</th>
<th>Std dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bedford</td>
<td>2.275</td>
<td>.399</td>
</tr>
<tr>
<td>MCH</td>
<td>2.467</td>
<td>.292</td>
</tr>
<tr>
<td>SWAT</td>
<td>2.183</td>
<td>.254</td>
</tr>
<tr>
<td>NASA-TLX</td>
<td>2.200</td>
<td>.212</td>
</tr>
</tbody>
</table>

TABLE 2: Means and standard deviations of applicability indices for two scenarios (N=10)

All four workload measures were found to be more applicable to the measurement of workload in the air-to-ground scenario than the air-to-air scenario. However, the only measure which was found to be significantly more applicable to the air-to-ground scenario was the Bedford scale (p<.002).

The average applicability index assigned to the workload measures was also calculated for each segment of each scenario. The means and standard deviations of these applicability indices are illustrated in table 3 below.

| T-OFF. & CL. ATT. APP. & LAND. |
|------------------------------|------------------------------|
| Workload measure             | Mean Std dev | Mean Std dev | Mean Std dev |
| Bedford                       | 1.750 .677   | 2.350 .560   | 2.500 .333   |
| MCH                           | 1.733 .540   | 2.700 .292   | 2.933 .141   |
| SWAT                          | 1.633 .331   | 2.367 .367   | 2.567 .274   |
| NASA-TLX                      | 1.617 .416   | 2.700 .131   | 2.650 .183   |

T-OFF. & CL. = TAKE-OFF AND CLIMB
ATT. = ATTACK
APP. & LAND. = APPROACH AND LANDING

TABLE 3: Means and standard deviations of applicability indices for each segment of each scenario (N=10)

Within the air-to-air scenario, the NASA-TLX, the SWAT and the MCH scale were found to be significantly more applicable during the "attack" segment than during "take-off and climb" (p<.000, p<.001, p<.000 respectively). Within the air-to-ground scenario, the same scales were again found to be significantly more applicable during the "attack" segment than during the "approach and landing" segment (p<.000, p<.000, p<.001 respectively). Thus, within each flight scenario, a significant difference in applicability indices according to flight segment was observed.
Qualitative data: Additional comments made on questionnaire

Respondents suggested the following factors would have contributed to their workload during the scenarios:

- The accuracy of information given to the pilot.
- The thoroughness of the pre-flight briefing.
- The length of time spent on "alert" before the mission.
- Weather conditions and time of day/night.
- Fear in war-time situation.
- The amount and accuracy of enemy ground and air defence fire in a war-time situation.
- The performance of others, such as the wingman or ground controller.
- The pilot's personal perception of the prevailing situation.
- The adequacy of communications.
- An individual pilot's experience and skill.

**PHASE II: THE SUITABILITY OF THE SUBJECTIVE WORKLOAD MEASURES**

The average suitability index assigned by subjects to each measure was calculated and is illustrated in table 4 below.

<table>
<thead>
<tr>
<th>Workload measure</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bedford</td>
<td>7.719</td>
<td>1.414</td>
</tr>
<tr>
<td>MCH</td>
<td>7.250</td>
<td>1.197</td>
</tr>
<tr>
<td>SWAT</td>
<td>5.719</td>
<td>2.702</td>
</tr>
<tr>
<td>NASA-TLX</td>
<td>5.250</td>
<td>2.258</td>
</tr>
</tbody>
</table>

**TABLE 4: Means and standard deviations of suitability indices assigned to each workload measure (N=16)**

Statistical analysis revealed the Bedford scale to be significantly more suitable than the NASA-TLX (p<.001) and the SWAT (p<.002). The MCH scale was also significantly more suitable than the NASA-TLX (p<.004).

Reasons given for high ("*") or low ("•") suitability indices are shown below; the percentage of subjects having expressed these views is shown alongside.

1) Bedford scale:
- Straightforward 38%
- Gives hint of how workload is measured/described 13%
- Merely requires "yes" or "no" decisions 13%
- Decision-tree structure too complex 6%
- Final definitions too lengthy/confusing 19%
2) MCH scale:
- Merely requires "yes" or "no" decisions 25%
- Good, concise definitions 6%
- Decision-tree structure too complex 13%

3) SWAT:
- Straightforward 31%
- Clarity of headings aids rating process 6%
- Too complicated and lengthy 38%

4) NASA-TLX:
- Clarity of headings aids rating process 13%
- Too lengthy and too many headings (dimensions) 44%

Suitability of subjective measures for different manoeuvres
- The suitability of the subjective measures considered in the study is very limited; none of the measures would be suitable for circuit flying, formation flying, weapon arming, low-level flying, attack or visual approach in unknown territory. 44%
- The unidimensional measures would be suitable for more manoeuvres than multidimensional measures. 19%

Post-task administration of measures
- Workload ratings given following a flight would be affected by the pilot's memory. 75%
- Providing the pilot is debriefed soon after landing, post-task ratings would not be affected by memory. 25%

The use of a keypad in-flight collection of workload ratings
- Keypad, or similar device, would not be more suitable than subjective measures. 44%
- Keypad, or similar device, would be more suitable than subjective measures. 44%
- Keypad could be more suitable than subjective measures when used in conjunction with a post-flight debrief. 12%
SUMMARY OF PRINCIPAL FINDINGS

1) The Bedford scale was found to be the most suitable for implementation during an orbit in a fast jet.

2) The unidimensional measures were generally found to be more suitable for implementation during fixed wing, military flight than the multidimensional measures.

3) The MCH scale was found to be the most applicable to the measurement of workload in the two scenarios.

4) The Bedford scale was found to be significantly more applicable to the assessment of workload in the air-to-ground scenario than in the air-to-air scenario.

5) The applicability of the NASA-TLX, the SWAT and the MCH scale was found to differ significantly between flight segments.
DISCUSSION

PHASE I: THE APPLICABILITY OF THE SUBJECTIVE WORKLOAD MEASURES

The results suggest that a measure which is applicable to a particular flight scenario may not be applicable to another. Furthermore, a measure which is applicable to a particular flight segment will not necessarily be applicable to another. The Bedford scale would seem to address fundamental factors which differ in their applicability across the two scenarios, but do not differentiate between segments. The other measures appear to address more specific factors which may be relevant at some point during both scenarios; however, when the applicability of these factors to different segments is examined, differences become apparent.

It should be appreciated that the fact that the MCH scale was found to be the most applicable measure does not imply that it addresses all relevant factors, but merely that those which it does address are indeed relevant. Respondents were also requested to suggest any additional factors which they felt would have contributed to their workload but which were not adequately addressed by the workload measures. Comments made by respondents suggest that there are some factors particularly relevant to the aviation environment which are not adequately addressed by any of the measures, such as the performance of others (e.g. wingman or ground controller), weather conditions, accuracy of information given to pilot and adequacy of communications. It is interesting to note that the majority of the additional contributing factors are "external" to the individual. This serves to emphasise the fact that a pilot does not perform as an isolated unit but as part of an interdependent team; the workload of the pilot is therefore, to a certain extent, dependent on how well the components of this team interact.

Many of the additional factors suggested are unique to the aviation environment, and may not be relevant to another. This suggests that, in order to provide a "complete" and accurate picture of workload in an operational environment, workload measures should ideally be adjusted to meet the needs of specific fields.

PHASE II: THE SUITABILITY OF THE SUBJECTIVE WORKLOAD MEASURES

It is evident from the results that the two unidimensional measures were, on average, considered to be more suitable than the multidimensional measures. This is probably largely due to the decision-tree structure of the unidimensional measures; the respondent is merely required to decide between a "yes" or "no" answer at each stage of the tree. Furthermore, the answer given at each stage immediately serves to eliminate some options, thus "narrowing down" the selection of final workload ratings. The respondent is therefore "guided" towards a final workload rating.
It should be noted that the suitability indices were assigned with a view to verbal presentation of the measures. A high average suitability index therefore suggests a measure to be suitable for implementation in single-seat combat aircraft, since radio communication could be maintained with a researcher on the ground. Judging from the average suitability indices, the two unidimensional measures could be implemented in this fashion. However, it should also be noted that the suitability indices were assigned with a view to the implementation of the measures during an orbit. Although none of the measures were considered to be totally unsuitable for implementation during an orbit, comments from participants during the post-task interview did indeed suggest the suitability of the measures for implementation during other manoeuvres, such as low-level or formation flying, to be limited.

The fact that the Bedford scale was found to be the most suitable of the four measures, and that many participants found it to be the most "straightforward", is in accordance with the findings of Lidderdale (op cit.): "The use of the 'Bedford Scale' was found to be easily understood and gained ready acceptance from the aircrew on the trial. Despite initial doubts, it was found that both pilots and navigators were able to give assessments of workload during flight, based on the scale, even under the most demanding flight conditions" (page 35).

When asked to comment on post-task administration of the subjective measures, the majority of participants were of the opinion that workload ratings given following a flight would indeed be affected by the pilots' memory - and also by the pilots' background, training and ego. It was generally felt that post-task ratings obtained from test-pilots should not be too biased since these pilots are trained to critically evaluate their own performance as well as that of the aircraft. The remaining participants felt that, providing the pilot is debriefed soon after landing, workload ratings given following the flight would not be affected by memory. In addition, if the pilot is fully informed prior to the flight of the ratings required, the provision of fairly accurate and reliable post-task ratings should not be too difficult.

A considerable number of participants were emphatic in their disapproval of the use of a keypad, or similar device, for recording workload ratings during flight. Some felt that it would be difficult to time such ratings appropriately whilst others expressed concern that such techniques ignore the reasons behind a particular rating. Conversely, many participants felt such techniques to be more suitable than subjective measures in providing an instantaneous measure of workload; "keying" a workload rating into a keypad was not thought to interfere with the flying task since "pilots are used to pressing buttons anyway". The remaining participants felt that the use of a keypad could be more suitable than subjective measures when used in conjunction with a post-flight debrief. The debrief could be used to provide more information about the reasons behind particular workload ratings.
CONCLUSIONS

Perusal of the results suggests the two unidimensional scales to be more suitable for implementation during fixed wing military flight than the multidimensional measures. A clear demarcation is evident between the higher suitability indices assigned to the unidimensional scales and those assigned to the multidimensional measures. Of the two unidimensional scales, the Bedford scale was assigned the highest average suitability index, followed closely by the MCH scale.

The MCH scale was also found to be the most applicable to the measurement of workload in the two scenarios, whereas the Bedford scale was found to be the third most applicable of the four measures. It would therefore appear that, of the four workload measures assessed, the MCH scale fared best on the criteria of both applicability and suitability. However, it should not be forgotten that the applicability and suitability of all the measures was found to be highly dependent on the task situation, i.e. the particular manoeuvre, flight segment or flight scenario being performed.

The results of the current study clearly indicate the importance of the criteria of applicability and suitability when assessing workload during military flight, or indeed any operational environment. The dearth of studies in the literature regarding these criteria emphasises the urgent need for further research into the area.

The findings indicate that differences in the applicability of various workload measures do indeed exist. Further research into the incidence of such differences is necessary; the situations during which the various workload measures are most applicable can then be determined. Moreover, further research is necessary to determine during which situations the various assessment measures can best be implemented.
ACKNOWLEDGEMENTS

The authors would like to express their gratitude to the staff of the ITPS, particularly Shawn Coyle, for their keen interest and assistance throughout this study.

We would also like to thank all those pilots who participated in the study.
REFERENCES


Ergonomics, 29 (11), 1385-1398.

Charles E. Merrill Publishing Co.: Ohio, USA.

global mental workload measurement applications. Proceedings of the 
Human Factors Society 27th Annual Meeting.
APPENDIX A

The Bedford and MCH scales and dimensions of the SWAT and NASA-TLX
level | descriptor
--- | ---

**Mental Effort Load**

1 | Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

2 | Moderate conscious mental effort required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

3 | Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

**Time Load**

1 | Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

2 | Occasionally have spare time. Interruptions or overlap among activities occur frequently.

3 | Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

**Psychological Stress Load**

1 | Little confusion, risk, frustration or anxiety exists and can be easily accommodated.

2 | Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload. Significant compensation is required to maintain adequate performance.

3 | High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self control required.

*SWAT scale descriptors*
<table>
<thead>
<tr>
<th>Title</th>
<th>Endpoints</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>MENTAL DEMAND</td>
<td>Low/High</td>
<td>How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?</td>
</tr>
<tr>
<td>PHYSICAL DEMAND</td>
<td>Low/High</td>
<td>How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?</td>
</tr>
<tr>
<td>TEMPORAL DEMAND</td>
<td>Low/High</td>
<td>How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?</td>
</tr>
<tr>
<td>PERFORMANCE</td>
<td>Perfect/</td>
<td>How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?</td>
</tr>
<tr>
<td></td>
<td>Failure</td>
<td></td>
</tr>
<tr>
<td>EFFORT</td>
<td>Low/High</td>
<td>How hard did you have to work (mentally and physically) to accomplish your level of performance?</td>
</tr>
<tr>
<td>FRUSTRATION LEVEL</td>
<td>Low/High</td>
<td>How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?</td>
</tr>
</tbody>
</table>

*NASA-TLX descriptors*
DECISION TREE

WORKLOAD DESCRIPTION

RATING

<table>
<thead>
<tr>
<th>Workload Insignificant</th>
<th>WL 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workload Low</td>
<td>WL 2</td>
</tr>
<tr>
<td>Enough spare capacity for all desirable additional tasks</td>
<td>WL 3</td>
</tr>
<tr>
<td>Insufficient spare capacity for easy attention to additional tasks</td>
<td>WL 4</td>
</tr>
<tr>
<td>Reduced spare capacity; additional tasks cannot be given the desired amount of attention</td>
<td>WL 5</td>
</tr>
<tr>
<td>Little spare capacity; level of effort allows little attention to additional tasks</td>
<td>WL 6</td>
</tr>
<tr>
<td>Very little spare capacity but maintenance of effort in the primary tasks not in question</td>
<td>WL 7</td>
</tr>
<tr>
<td>Very high workload with almost no spare capacity, difficulty in maintaining level of effort</td>
<td>WL 8</td>
</tr>
<tr>
<td>Extremely high workload, no spare capacity, serious doubts as to ability to maintain level of effort</td>
<td>WL 9</td>
</tr>
<tr>
<td>Task abandoned, pilot unable to apply sufficient effort</td>
<td>WL 10</td>
</tr>
</tbody>
</table>

The Bedford Scale
### Rating Scale

<table>
<thead>
<tr>
<th>Difficulty Level</th>
<th>Operator Demand Level</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Easy, Highly Desirable</td>
<td>Operator mental effort is minimal and desired performance is easily attainable</td>
<td>1</td>
</tr>
<tr>
<td>Easy, Desirable</td>
<td>Operator mental effort is low and desired performance is attainable</td>
<td>2</td>
</tr>
<tr>
<td>Fair, Mild Difficulty</td>
<td>Acceptable operator mental effort is required to attain adequate system performance</td>
<td>3</td>
</tr>
<tr>
<td>Minor but Annoying Difficulty</td>
<td>Moderately high operator mental effort is required to attain adequate system performance</td>
<td>4</td>
</tr>
<tr>
<td>Moderately Objectionable Difficulty</td>
<td>High operator mental effort is required to attain adequate system performance</td>
<td>5</td>
</tr>
<tr>
<td>Very Objectionable but Tolerable Difficulty</td>
<td>Maximum operator mental effort is required to attain adequate system performance</td>
<td>6</td>
</tr>
<tr>
<td>Major Difficulty</td>
<td>Maximum operator mental effort is required to bring errors to moderate level</td>
<td>7</td>
</tr>
<tr>
<td>Major Difficulty</td>
<td>Maximum operator mental effort is required to avoid large or numerous errors</td>
<td>8</td>
</tr>
<tr>
<td>Major Difficulty</td>
<td>Intense operator mental effort is required to accomplish task, but frequent or numerous errors persist</td>
<td>9</td>
</tr>
<tr>
<td>Impossible</td>
<td>Instructed task cannot be accomplished reliably</td>
<td>10</td>
</tr>
</tbody>
</table>
APPENDIX B

Applicability questionnaire, flight scenarios and covering letters
June 1990

Dear Sir,

The Cranfield Institute of Technology is currently involved in a large-scale project initiated by the Royal Aircraft Establishment at Farnborough concerning the assessment of pilot workload. Part of this research aims to investigate the applicability of various subjective workload assessment techniques. Numerous subjective techniques have been developed for the assessment of pilot workload; each is based on a number of factors seen to make up workload. However, the type and number of factors vary according to technique. The current research breaks down each of four techniques into their component factors so that the applicability of each factor, and thus of the overall technique, may be assessed.

As you will appreciate, the assessment of workload is very individual, being influenced by the pilots flying experience. For this reason, it is important to request the assistance of as many fast jet pilots as possible. Attached to this letter is a questionnaire devised specifically for this project. Please retain this for your reference. In addition, each of the enclosed envelopes contains a copy of this questionnaire, a covering letter and a freepost return envelope. I would be very grateful if you would distribute the enclosed envelopes to ten pilots at your Squadron.

If you require further information please do not hesitate to contact me at the above address. Many thanks.

Yours sincerely,

Miss A. Markou BSc (Hons)
June 1990

Dear Sir,

The Cranfield Institute of Technology is currently involved in a large-scale project initiated by the Royal Aircraft Establishment at Farnborough concerning the assessment of pilot workload. Part of this research aims to investigate the applicability of various subjective workload assessment techniques. Numerous subjective techniques have been developed for the assessment of pilot workload; each is based on a number of factors seen to make up workload. The current research breaks down each of four techniques into their component factors so that the applicability of each factor, and thus of the overall technique, may be assessed.

As you will appreciate, the assessment of workload is very individual, being influenced by the pilots flying experience. For this reason, it is important to request the assistance of as many fast jet pilots as possible.

Attached to this letter is a questionnaire devised specifically for this study. Completion of this questionnaire is purely voluntary. Moreover, your name is not required, only some brief details regarding your flying experience. The questionnaire requires you to rate the applicability of various assessment techniques in describing the workload imposed during two flight scenarios. Of course, it is recognised that such a retrospective evaluation is limited and cannot replace an evaluation following the actual performance of the flight scenarios. Consequently, the current stage of the study is exploratory and paves the way for the actual in-flight investigation of the techniques.

If you are willing to participate in this study, please read the instructions on the first page of the questionnaire. Please return your completed questionnaire to Andie Markou in the freepost envelope provided as soon as possible. If you require further information do not hesitate to contact me at the above address. Many thanks.

Yours sincerely,

Miss A. Markou BSc (Hons)
SUBJECTIVE WORKLOAD ASSESSMENT

APPLICABILITY QUESTIONNAIRE
Instructions

An air-to-air (counter-air) and an air-to-ground (close air support) flight scenario are illustrated in the blue leaflet attached to this questionnaire. Detailed descriptions of the whole scenarios have been omitted as these would have been unnecessarily lengthy and time-consuming. Instead, various segments of the scenarios have been isolated and are described in detail.

Please begin by examining the segments of Flight Scenario 1 (air-to-air) carefully; imagine you have been asked to fly these segments in a fixed wing fighter aircraft. No further specifications will be made as to aircraft type; you are required to use your own personal experience and judgement. You may imagine executing the flight segments in a fighter aircraft which you have recently flown or in which you feel experienced.

Next turn to the first page of this questionnaire.

You are required to rate the applicability of 4 workload assessment techniques in describing the workload imposed during each of the flight scenario segments. The dimensions or factors on which each technique is based are presented in tabular form. You are required to indicate how applicable each factor would be in the workload experienced during each segment by circling any one of the following responses:

VA = Very Applicable
FA = Fairly Applicable
WA = Not Applicable.

Please begin by assessing the applicability of Assessment Technique 1 and then repeating this process with Techniques 2, 3 and 4.

When you have completed rating the applicability of the 4 assessment techniques for Flight Scenario 1, please repeat the process for Flight Scenario 2 (air-to-ground).

Any additional comments you would like to make regarding the various assessment techniques should be made in the spaces provided.

Thank-you for your co-operation and assistance.
**FLIGHT SCENARIO 1 (COUNTER-AIR)**

**ASSESSMENT TECHNIQUE 1**

<table>
<thead>
<tr>
<th>WORKLOAD DIMENSIONS/FACTORS ON WHICH TECHNIQUE IS BASED</th>
<th>FLIGHT SEGMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TAKE-OFF &amp; CLIMB</td>
</tr>
<tr>
<td>Amount of mental and perceptual activity required (eg. calculating, thinking, remembering, searching).</td>
<td>VA</td>
</tr>
<tr>
<td>Amount of physical activity required (eg. pushing, pulling, turning).</td>
<td>VA</td>
</tr>
<tr>
<td>Amount of time pressure felt due to the rate or pace at which the task or task elements occurred.</td>
<td>VA</td>
</tr>
<tr>
<td>Satisfaction with own performance: success in meeting goals of task.</td>
<td>VA</td>
</tr>
<tr>
<td>Effort (mental and physical) necessary to accomplish level of performance.</td>
<td>VA</td>
</tr>
<tr>
<td>Feelings of insecurity, irritation, discouragement, stress and annoyance experienced during task.</td>
<td>VA</td>
</tr>
</tbody>
</table>

In your opinion, are there any other factors which would have contributed to your workload during these segments? If so, please list them below, and include a very brief explanation of why you feel these factors are applicable:

____________________________________________________________________________________

____________________________________________________________________________________

____________________________________________________________________________________

____________________________________________________________________________________

Further comments:

____________________________________________________________________________________

____________________________________________________________________________________

____________________________________________________________________________________
ASSESSMENT TECHNIQUE 2

<table>
<thead>
<tr>
<th>WORKLOAD DIMENSIONS/FACTORs ON WHICH TECHNIQUE IS BASED</th>
<th>FLIGHT SEGMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TAKE-OFF &amp; CLimb</td>
</tr>
<tr>
<td>The amount of conscious mental effort, concentration and attention required.</td>
<td>VA FA NA</td>
</tr>
<tr>
<td>The amount of spare time available, interruptions and overlap among activities.</td>
<td>VA FA NA</td>
</tr>
<tr>
<td>Feelings of confusion, frustration, risk or anxiety experienced and the amount of compensation required to maintain adequate performance.</td>
<td>VA FA NA</td>
</tr>
</tbody>
</table>

In your opinion, are there any other factors which would have contributed to your workload during these segments? If so, please list them below, and include a very brief explanation of why you feel these factors are applicable:

Further comments:
### ASSESSMENT TECHNIQUE 3

#### WORKLOAD DIMENSIONS/FACTORS ON WHICH TECHNIQUE IS BASED

<table>
<thead>
<tr>
<th>Factor</th>
<th>TAKE-OFF &amp; CLimb</th>
<th>ATTACK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whether errors are small/large, inconsequential/consequential, infrequent/frequent.</td>
<td>VA FA NA</td>
<td>VA FA NA</td>
</tr>
<tr>
<td>The difficulty of the task.</td>
<td>VA FA NA</td>
<td>VA FA NA</td>
</tr>
<tr>
<td>The degree of operator mental effort required to attain desired performance.</td>
<td>VA FA NA</td>
<td>VA FA NA</td>
</tr>
</tbody>
</table>

In your opinion, are there any other factors which would have contributed to your workload during these segments? If so, please list them below, and include a very brief explanation of why you feel these factors are applicable:

Further comments:
ASSESSMENT TECHNIQUE 4

<table>
<thead>
<tr>
<th>WORKLOAD DIMENSIONS/FACTORS ON WHICH TECHNIQUE IS BASED</th>
<th>FLIGHT SEGMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TAKE-OFF &amp; CLIMB</td>
</tr>
<tr>
<td>Possibility of completing the task.</td>
<td>VA</td>
</tr>
<tr>
<td>The amount of spare capacity available for additional tasks.</td>
<td>VA</td>
</tr>
</tbody>
</table>

In your opinion, are there any other factors which would have contributed to your workload during these segments? If so, please list them below, and include a very brief explanation of why you feel these factors are applicable:

Further comments:

- 5 -
FLIGHT SCENARIO 2 (CLOSE AIR SUPPORT)

ASSESSMENT TECHNIQUE 1

<table>
<thead>
<tr>
<th>WORKLOAD DIMENSIONS/FACTORS ON WHICH TECHNIQUE IS BASED</th>
<th>FLIGHT SEGMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ATTACK</td>
</tr>
<tr>
<td>Amount of mental and perceptual activity required (eg. calculating, thinking, remembering, searching).</td>
<td>VA FA NA</td>
</tr>
<tr>
<td>Amount of physical activity required (eg. pushing, pulling, turning).</td>
<td>VA FA NA</td>
</tr>
<tr>
<td>Amount of time pressure felt due to the rate or pace at which the task or task elements occurred.</td>
<td>VA FA NA</td>
</tr>
<tr>
<td>Satisfaction with own performance: success in meeting goals of task.</td>
<td>VA FA NA</td>
</tr>
<tr>
<td>Effort (mental and physical) necessary to accomplish level of performance.</td>
<td>VA FA NA</td>
</tr>
<tr>
<td>Feelings of insecurity, irritation, discouragement, stress and annoyance experienced during task.</td>
<td>VA FA NA</td>
</tr>
</tbody>
</table>

In your opinion, are there any other factors which would have contributed to your workload during these segments? If so, please list them below, and include a very brief explanation of why you feel these factors are applicable:

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

Further comments:

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

- 6 -
**ASSESSMENT TECHNIQUE 2**

<table>
<thead>
<tr>
<th>WORKLOAD DIMENSIONS/FACTORS ON WHICH TECHNIQUE IS BASED</th>
<th>FLIGHT SEGMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>The amount of conscious mental effort, concentration and attention required.</td>
<td>ATTACK</td>
</tr>
<tr>
<td></td>
<td>VA</td>
</tr>
<tr>
<td>The amount of spare time available. Interruptions and overlap among activities.</td>
<td>VA</td>
</tr>
<tr>
<td>Feelings of confusion, frustration, risk or anxiety experienced and the amount of compensation required to maintain adequate performance.</td>
<td>VA</td>
</tr>
</tbody>
</table>

In your opinion, are there any other factors which would have contributed to your workload during these segments? If so, please list them below, and include a very brief explanation of why you feel these factors are applicable:

Further comments:

---

- 7 -
### ASSESSMENT TECHNIQUE 3

<table>
<thead>
<tr>
<th>WORKLOAD DIMENSIONS/FACTORS ON WHICH TECHNIQUE IS BASED</th>
<th>FLIGHT SEGMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ATTACK</td>
</tr>
<tr>
<td>Whether errors are small/large, inconsequential/consequential, infrequent/frequent.</td>
<td>VA</td>
</tr>
<tr>
<td>The difficulty of the task.</td>
<td>VA</td>
</tr>
<tr>
<td>The degree of operator mental effort required to attain desired performance.</td>
<td>VA</td>
</tr>
</tbody>
</table>

In your opinion, are there any other factors which would have contributed to your workload during these segments? If so, please list them below, and include a very brief explanation of why you feel these factors are applicable:

[Further comments]

---

Further comments:
ASSESSMENT TECHNIQUE 4

<table>
<thead>
<tr>
<th>WORKLOAD DIMENSIONS/FACTORS ON WHICH TECHNIQUE IS BASED</th>
<th>FLIGHT SEGMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Possibility of completing the task.</td>
<td>ATTACK APP. &amp; LANDING</td>
</tr>
<tr>
<td>The amount of spare capacity available for additional tasks.</td>
<td>VA FA NA VA FA NA</td>
</tr>
</tbody>
</table>

In your opinion, are there any other factors which would have contributed to your workload during these segments? If so, please list them below, and include a very brief explanation of why you feel these factors are applicable:

Further comments:
Finally, could you please specify which aircraft you had in mind when visualising the flight segments?:

________________________________________________________________________

Approximately how much experience (flying hours) have you had in flying this/these aircraft?:

________________________________________________________________________
SUBJECTIVE WORKLOAD ASSESSMENT

FLIGHT SCENARIOS

1) COUNTER-AIR (AIR-TO-AIR) SCENARIO

Air offense
- Fighter sweep
- Fighter escort
- Combat air patrol

Air defense
- Point intercept
- Area defense
- CAP loiter

Figure I: Counter-air (air-to-air) scenario

a) Take-off and climb

- Release brakes
- Make GO/NO GO - decision on speed checks
- Perform rotation and lift-off
- Maintain directional control
- Establish climb attitude or angle
- Call airbourne
- Reconfigure aircraft (gear, flaps)
- Establish climb speed
- Establish departure course
- Maintain climb speed/follow pitch steering command
- Contact control agency
b) **Attack**

Obtain target assignment instruction  
Identify/acquire/designate/evaluate target data on radar/tactical display  
Acknowledge assignment information and data to IC  
Establish and check intercept attack programming  
(Note/insert/read)  
Establish and check intercept flight profile data (speed, altitude)  
Establish and check command and control data (heading, break-away headings)  
Establish and check command and control data (frequency, brevity code R/T procedures)  
Establish and check armament data (selection, arming status)  
Check aircraft systems (fuel, oxygen, avionic status)  
Make GO/NO GO decision  
Obtain final order/clearance for (ID-pass, intervention) engagement  
Authenticate engagement order  
Recheck fuel, armament status  
Arm selected weapons  
Establish attack airspeed (combat speed)  
Acknowledge target bearing and range  
Take over intercept by "Judy"  
Engage automatic maneuvering system  
Monitor/follow maneuver commands  
Monitor aircraft steering  
Monitor weapons envelope/firing range; determine time to launch and time of commitment (target radar coverage time for update purposes)  
Consider weapon options; visualize weapons, delivery procedures, firing doctrine  
Recheck weapons ready and armed  
Monitor shoot lights, actuate trigger  
Confirm weapons release
2) CLOSE AIR SUPPORT (AIR-TO-GROUND) SCENARIO

![Diagram of close air support scenario]

- Release from Orbit
- Hold at Contact
- FAC Freq/Call Sign
- Mission Number

Target Info
IP position
Range/Bearing
Laser Coding

Figure II: Close-air-support (air-to-ground) scenario

a) Attack

Depart initial point on heading, airspeed, altitude with time reference
Navigate to target area
Maintain tactical formation
Maintain visual/sensor look-out
Employ countermeasures
Perform system monitoring (fuel)
Navigate to pop-up point
Establish slight climb at pop-up point
Acquire target with sensors (helmet mounted sight, radar, laser, IR, etc.)
Designate target (helmet mounted sight, radar, laser, IR, etc.)
Establish high "$g\"", level turn
Engage automated weapon delivery system
Monitor aircraft steering and performance
Prepare alternate delivery mode
Make decision to release weapon
Activate weapon release button
Confirm weapon release
Perform recovery
Obtain intelligence data and BDA
Monitor countermeasures operation
Make decision to egress or re-attack

b) Approach and landing

Review approach procedures
Monitor comm and IFF
Perform weather penetration complying with published procedures or controllers instructions
Acquire airbase visually or with sensors
Configure aircraft for landing
Accomplish landing checks
Perform manual or automated (ILS, GCA, ARA or AILA) final approach
Acquire landing area visually or with sensors
Obtain landing clearance
Determine desired touchdown point
Determine aircraft controllability (gross weight, hung ordance, battle damage, etc.)
Establish landing parameters manually or automated (airspeed, AOA, sink rate, flight path angle, ground track)
Transition from stabilized approach to flare
Perform touchdown
Maintain lateral spacing on runway/landing area
Slow aircraft to taxi speed
Perform turn-off of runway/landing area

***********************
APPENDIX C

Counterbalanced design for presentation of workload measures
ORDER OF PRESENTATION OF MEASURES

<table>
<thead>
<tr>
<th>SUBJECTS</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
</table>

Key: UNI = unidimensional, MULTI = multidimensional, Bed = Bedford scale, NASA = NASA-TLX

Counterbalanced design for presentation of workload measures
APPENDIX D

Revised SWAT and NASA-TLX dimensions used in current study
"Mental effort load: Conscious mental effort or concentration required; was it very little, moderate or extensive?"

"Time load: Spare time; did you often, occasionally or almost never have spare time?"

"Psychological stress load: Feelings of confusion, risk, frustration or anxiety; were they little, moderate or high?"

Questions posed to subjects during the SWAT rating ("event scoring") stage

"Mental demand: How much mental activity was required?"

"Physical demand: How much physical activity was required?"

"Temporal demand: How much time pressure did you feel?"

"Performance: How successful do you think you were in accomplishing the goals of the task?"

"Effort: How hard did you have to work, mentally and physically?"

"Frustration level: How insecure, discouraged, irritated, stressed and annoyed did you feel?"

Questions posed to subjects during the NASA-TLX rating ("magnitude of load") stage