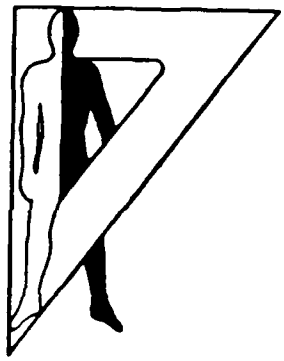
(4)



AD

Technical Note 7-88

AN ANNOTATED BIBLIOGRAPHY ON OPERATOR MENTAL WORKLOAD ASSESSMENT.

John K. Schmidt
Helen M. Nicewonger

August 1988
AMCMS Code 612716.H700011

U. S. ARMY HUMAN ENGINEERING LABORATORY
Aberdeen Proving Ground, Maryland

Destroy this report when no longer needed.
Do not return it to the originator.

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Use of trade names in this report does not constitute an official endorsement or approval of the use of such commercial products.

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1a REPORT SECURITY CLASSIFICATION Unclassified	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b DECLASSIFICATION / DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Note 7-88	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION Human Engineering Laboratory 6b OFFICE SYMBOL (If applicable) SLCHE	7a NAME OF MONITORING ORGANIZATION
6c ADDRESS (City, State, and ZIP Code) Aberdeen Proving Ground, Maryland 21005-5001	7b ADDRESS (City, State, and ZIP Code)
8a NAME OF FUNDING / SPONSORING ORGANIZATION 8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 6.27.16 / PROJECT NO. 1L162716AH70 / TASK NO. / WORK UNIT ACCESSION NO.

11 TITLE (Include Security Classification)
An Annotated Bibliography on Operator Mental Workload Assessment

12 PERSONAL AUTHOR(S) Schmidt, John K.
Nicewonger, Helen M.

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Final	FROM _____ TO _____	1988, August	103

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	workload, operator workload, mental workload, workload assessment
05	08		
23	02		

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

An annotated bibliography on operator mental workload assessment is provided with corresponding documentation to enhance its utility as a reference. The present work is to be viewed as an extension to and not a replacement for an earlier effort by Wierwille and Williges (1980), An Annotated Bibliography on Operator Mental Workload Assessment. For the most part, references published between the years 1980 and 1986 are included. Each of the 206 citations presented from the literature contain reference information as well as an abstract. All listings are indexed by both author and subject.

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT	21 ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL Doris S. Eanes	22b TELEPHONE (Include Area Code) 301-278-4478 22c OFFICE SYMBOL SLCHE-SS-IR

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE

AMCMS Code 612716.H700011 Technical Note 7-88

AN ANNOTATED BIBLIOGRAPHY ON OPERATOR MENTAL WORKLOAD ASSESSMENT

John K. Schmidt
Helen M. Nicewonger

August 1988

APPROVED: _____
JOHN D. WEISZ
Director
Human Engineering Laboratory

Approved for public release;
distribution is unlimited.

U.S. ARMY HUMAN ENGINEERING LABORATORY
Aberdeen Proving Ground, Maryland 21005-5001

ACKNOWLEDGMENTS

The compilers of this annotated bibliography would like to thank two members of the Aviation and Air Defense Division (AADD), Human Engineering Laboratory (HEL), for their clerical support of this project--Ms. Emilie Waddington and Ms. Priscilla Devonshire. The generation of this document would not have come to pass without their assistance. In addition, we would be remiss by not expressing our gratitude to Mr. Alan Poston, Aviation Team Leader, AADD, for his assistance. Finally, we would like to express our appreciation to Ms. Joyce Watlington, Services and Support Division, HEL, for her technical support.

CONTENTS

AN ANNOTATED BIBLIOGRAPHY ON OPERATOR MENTAL WORKLOAD ASSESSMENT

BACKGROUND

The integration of modern technology into military systems has greatly increased the cognitive demands placed on operators. It is contended that the amount of "mental workload" placed on a system operator is directly related to his performance and that excessive demands will degrade system effectiveness and adversely impact mission completion. In order to determine what requirements are placed on an operator, a number of assessment techniques have been employed.

However, many difficulties have been cited in measuring mental workload (Wierwille & Williges, 1980). First, there have been problems in finding agreement on a definition for mental workload. Next, there is little consensus as to what technique should be employed to assess mental workload. Finally, most methods tend to vary in their applicability in given situations to estimate mental workload. One theoretical rationale that accounts for this difficulty is that the construct of mental workload is multifaceted and that any attempt to assess it must have such an orientation and use an array of measurement techniques.

The recognized significance of having the capability to determine mental workload levels has prompted numerous research efforts in many fields of endeavor. In effect, its importance has created a diverse literature that is not readily accessible to all who may be involved. The Aviation and Air Defense Division of the Human Engineering Laboratory, Aberdeen Proving Ground, Maryland, sensing this need, has opted to provide an updated annotated bibliography to the human factors community. The present work is to be viewed as an extension to and not a replacement for the earlier effort of Wierwille and Williges (1980). The 206 references listed were primarily published between the years 1980 and 1986 with the exception of a few earlier works not found in Wierwille and Williges (1980).

SEARCH PROCESS

The search process for covering the available literature on mental workload estimation entailed the use of several sources. Specifically, books, scientific journals, technical reports, and conference proceedings were combed via computer inquiries of on-line literature data bases, hand library searches, cross-referencing, and direct contact with mental workload researchers to identify appropriate items to be included.

Computer Inquiries

Searches were conducted using key words associated with the mental workload literature. Key words used are listed in Table 1. The DIALOG System was used to access the National Technical Information Service (NTIS) holdings.

Table 1

Key Words

Workload with:
Performance
Primary task
Secondary task
Physiological
Subjective
Operator
Estimation
Pilot
Space mental capacity
Sensory overlaod
Residual capacity
Residual attention
Perceptual loading
Time-sharing
Mental load
Attentional demands

Hand Searches

Manual extractions of sources were made to supplement the computerized searches. One important reason for conducting this operation is that much of the literature from technical conferences is not listed on automated literature data bases. The following sources were reviewed:

Annual Review of Psychology

Applied Ergonomics

Aviation, Space & Environmental Medicine

Ergonomics

Human Factors

The Institute of Electrical and Electronics Engineers (IEEE) Conference Proceedings

Human Factors Society Conference Proceedings

Journal of Experimental Psychology

Perceptual and Motor Skills

Psychological Bulletin

Psychological Review

Direct Contact

Noted members of the mental workload research community were contacted; and copies of their recent bibliographies, solicited.

SELECTION CRITERIA

The large number of documents identified through the various search processes made it necessary to establish criteria for limiting the number of references to be included. Generally, the topic had to be directly related to workload and its assessment; the population and application made had to fall within the realm of human factors type research; and the reference had to exist in print.

INDEXING PROCEDURE

Citations are listed alphabetically and are assigned a distinct identification number (001-206). Two indexes using these identification numbers were constructed. The first is a complete listing of all authors, and the second is by subject and subtopical areas. The authors hope that the index will increase the utility of the document.

ANNOTATED BIBLIOGRAPHY

001 Ackerman, P. L. (1984). Multivariate evaluation of workload, resources, and performance in dual task data. Proceedings of the Human Factors Society 28th Annual Meeting, 1, 210-214.

The problems of evaluating dual and complex task experiments are discussed. Two methods are proposed that provide a metric for measuring and comparing performance and workload in such experiments. The first method is based on derivations of effect size measures, which provide a useful descriptive and a valid inferential tool for dealing with single task, dual task, and multiple response measure situations. The second method involves a new approach to empirical derivation of performance-resource functions. The method is straightforward and testable. Examples and illustrations are presented.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 1, p. 210. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

002 Ackerman, P. L., Schneider, W., & Wickens, C. D. (1984). Deciding the existence of a time-sharing ability: A combined methodological and theoretical approach. Human Factors, 26, 71-82.

Experimental and statistical methods for examining individual differences in dual-task performance and time-sharing ability are reviewed and criticized. Previous data and analysis procedures are generally inadequate to evaluate a time-sharing ability. Errors resulting from unsophisticated use of correlational and factor analytic procedures are described. Four previous studies that concern time-sharing are considered in detail. The nature of task selection, scoring methods, and control of practice and reliability issues are discussed. Based on a reanalysis of available data, a time-sharing ability is not rejected. Simulation, incorporation of theory in planning models, and crucial tests of the hypotheses are proposed as methods for assessing the time-sharing ability.

(From Human Factors, 1984, 26, p. 71. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

003 Ackerman, P. L., & Wickens, C. D. (1982). Methodology and the use of dual and complex-task paradigms in human factors research. Proceedings of the Human Factors Society 26th Annual Meeting, pp. 354-358.

Two problem situations are discussed with respect to the multiple task literature, one stemming from statistical artifacts derived by inappropriate data-analytic procedures and the other caused by a failure to utilize specific models which map theory into data-level descriptions. These issues are discussed

as they apply to individual differences research and to investigations of cognitive processes, resources, and systems in multiple task performance. Examples and illustrations of these issues are provided.

(From Proceedings of the Human Factors Society 26th Annual Meeting, 1982, p. 354. Copyright 1982 by the Human Factors Society, Inc. Reprinted by permission.)

004 Acosta, E., Jr., & Dickman, J. D. (1984). Information processing in the visual periphery: Influence of mental workload. Proceedings of the Human Factors Society 28th Annual Meeting, 2, 737-739.

Two experiments examined the effects of the mental workload of various foveal task requirements on the identification of briefly presented symbols to the visual periphery. Results are discussed in light of two models (tunnel vision and constant decrement) that describe the relationship between mental workload and information processing in the visual periphery.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 2, p. 737. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

005 Acton, W. H., Crabtree, M. S., Simons, J. C., Gomer, F. E., & Eckel, J. S. (1983). Quantification of crew workload imposed by communications-related tasks in commercial transport aircraft. Proceedings of the Human Factors Society 27th Annual Meeting, 1, 239-243.

Information theoretic analysis and subjective paired-comparison and task ranking techniques were employed in order to scale the workload of 20 communications-related tasks frequently performed by the captain and first officer of transport category aircraft. Tasks were drawn from taped conversations between aircraft and air traffic controllers (ATC). Twenty crewmembers performed subjective message comparisons and task rankings on the basis of workload. Information theoretic results indicated a broad range of task difficulty levels, and substantial differences between captain and first officer workload levels. Preliminary subjective data tended to corroborate these results. A hybrid scale reflecting the results of both the analytical and the subjective techniques is currently being developed. The findings will be used to select representative sets of communications for use in high fidelity simulation.

(From Proceedings of the Human Factors Society 27th Annual Meeting, 1983, 1, p. 239. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

006 Antin, J. F., & Wierwille, W. W. (1984). Instantaneous measures of mental workload: An initial investigation. Proceedings of the Human Factors Society 28th Annual Meeting, 1, 6-10.

Many researchers have implicitly or explicitly averaged measures of mental workload over a given run or period of time. This process tends to mask specific features (e.g. peaks) of interest in the flow of instantaneous mental workload

(IMWL). A study was conducted in which thirty subjects performed computer tasks which quantifiably varied in difficulty during a run. These tasks emphasized perceptual, mediational, psychomotor, and storage and retrieval from short-term memory processes. Data were gathered on the following candidate measures of momentary load: instantaneous primary task performance, instantaneous Michon tapping and time estimation secondary task performance, instantaneous pulse and respiration rates, and two types of online subjective opinion. Data were short-term averaged and used to develop regression models to evaluate the ability of the measures to track IMWL. Primary task performance (response time for functional subtasks) and both forms of online subjective opinion measures showed great promise as measures of IMWL.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 1, p. 6. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

007 Arbak, C. J., Shew, R. L., & Simons, J. C. (1984). The use of reflective SWAT for workload assessment. Proceedings of the Human Factors Society 28th Annual Meeting, 2, 959-962.

The Subjective Workload Assessment Technique (SWAT) proposed by Reid et al. (1981) is a general measure of operator workload that is based on ratings of time, effort, and stress. In practice, SWAT has typically been used to assess workload at particular points during a task; however, it is also of practical use to estimate workload through an operator's reflections on their mission performance. The present effort depicts two examples of a reflective use of SWAT in combination with an interview to determine workload. From them, methodological conclusions are drawn and their significance is discussed.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 2, p. 959. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

008 Asiala, C. F., Loy, S. L., Welde, W. L., Bull, R. F., & Fitzgerald, J. A. (1981). Pilot workload assessment (AFAMRL TR-80-109). Wright-Patterson Air Force Base, OH: Air Force Aerospace Medical Research Laboratory. (DTIC No. AD-B058 801)

This study was conducted to perform a definitive assessment of a single-seat pilot's workload. The specific research objectives were to provide: (1) a methodology which can be used to predict effectively the workload incurred by a pilot in an advanced cockpit, and to determine the most effective instrument/display configuration and information format for minimizing pilot workloads; and (2) a data base that can be used for comparison of current to advanced single-seat cockpit concepts. The tactical role of the A-7D aircraft was examined and mission segments which created high pilot workload were identified. These were then subjected to a functions and task allocation analysis. Field testing techniques were proceduralized and data on pilot performance were acquired from A-7D operational bases. The field data were transformed into quantitative descriptors for use in computer-based integrated assessment techniques.

009 Bateman, R. P., Acton, W. H., & Crabtree, M. S. (1984). Workload and performance: Orthogonal measures. Proceedings of the Human Factors Society 28th Annual Meeting, 2, 678-679.

During an attempt to corroborate historical error data with recently gathered data on subjective workload values, it was noted that the data not only failed to show any degree of correlation, but gave strong evidence to support arguments that the workload and error measures are orthogonal (totally independent). This paper discusses several explanations of the observed phenomenon. A hypothesis is advanced that error free performance is strongly dependent upon task procedure construction, while workload is dependent upon subtask difficulty and situational stresses.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 2, p. 678. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

010 Battiste, V., & Hart, S. G. (1985). Predicted versus experienced workload and performance on a supervisory control task. Proceedings of the Third Symposium on Aviation Psychology, pp. 255-262.

Some of the stated goals in workload research are to provide information about operator workload in existing systems, predict the impact of modifications of existing systems, to provide designers with an accurate estimate of the expected workload of new systems at the inception of the design stage. This was the second in a series of studies conducted with a multi-task simulation of a supervisory control system. The operators' task was to complete a number of task elements (represented by different symbols), by selecting a target task and entering the desired command. Task difficulty and experienced workload were varied by manipulating the number of elements per task, the number of tasks, task schedule, and availability of task elements for performance. The goal was to investigate operators' abilities to predict the workload and performance impact of unfamiliar task features and configurations from their basic knowledge of the system and specific information provided before each scenario. Significant differences in performance and workload were found as a function of all experimental manipulations. In addition, different relationships between workload predictions and workload ratings were found due to the similarity of the task modifications to familiar levels and to task complexity.

011 Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. Psychological Bulletin, 91, 276-292.

A physiological measure of processing load or "mental effort" required to perform a cognitive task should accurately reflect within-task, between-task, and between-individual variations in processing demands. This article reviews all available experimental data and concludes that the task-evoked pupillary response fulfills these criteria. Alternative explanations are considered and rejected. Some implications for neurophysiological and cognitive theories of processing resources are discussed.

012 Bellenkes, A. H. (1984). Dual-task timesharing using a projected attitude display (Malcolm Horizon) (NAMRL TR-1310). Pensacola, FL: Naval Aerospace Medical Research Laboratory. (DTIC No. AD-A150 789)

In order to optimize timesharing and facilitate performance in the high workload environment of the modern cockpit, a Peripheral Vision Horizon Device (PVHD) has been developed which can present aircraft attitude data to the visual periphery; an area reported to be highly sensitive to the perception of information regarding orientation in space. A great deal of subjective evidence gathered from simulator and operational test flights has lent support to the efficacy of this device in improving performance. However, this capability has yet to be verified by controlled laboratory testing. Two horizon sizes were evaluated; one with dimensions similar to that found in an aircraft instrument panel and the other extending out to the visual periphery. The objective of this study was to determine whether dual-task performance could be improved by using the large projected horizon vs. a more conventional short horizon. The findings indicated that the PVHD allowed subjects to perform the foveated mental arithmetic task while simultaneously controlling the orientation of the horizon. PVHD root mean square (RMS) error, and mental arithmetic speed/accuracy data were found to be superior when subjects used the extended vs. the short horizon for tracking. These findings suggest that the PVHD permitted individuals to process the two sets of visual information in parallel, thereby improving performance on both.

013 Berg, S. L., & Sheridan, T. B. (1984). Measuring workload differences between short-term memory and long-term memory scenarios in a simulated flight environment. Proceedings of the 20th Annual Conference on Manual Control, NASA CP-2341, 1, 397-416.

Four highly experienced Air Force pilots each flew four simulated flight scenarios. Two scenarios required a great deal of aircraft maneuvering. The other two scenarios involved less maneuvering, but required remembering a number of items. All scenarios were designed to be equally challenging. Pilot's Subjective Ratings for Activity-level, Complexity, Difficulty, Stress, and Workload were higher for the maneuvering scenarios than the memory scenarios. At a moderate workload level, keeping the pilots active resulted in better aircraft control. When required to monitor and remember items, aircraft control tended to decrease. Pilots tended to weigh information about the spatial positioning and performance of their aircraft more heavily than other items.

014 Berg, S. L., & Sheridan, T. B. (1985). Effect of time span and task load on pilot mental workload (NASA CR-177388). Washington, DC: National Aeronautics and Space Administration.

Two sets of simulations are described that were designed to examine how a pilot's mental workload would be affected by continuous manual-control activity versus discrete mental tasks that included the length of time between receiving an assignment and executing it. A fixed-base flight simulator was used that consisted of a control box (joystick, throttle, switches for operating electronic and mechanical systems) and a high resolution CRT. Aircraft dynamics were modeled on a Lockheed Jetstar business jet. The CRT display consisted of a forward "out-the-window" perspective view and a cockpit instrument/indicator.

The first experiment evaluated two types of measures: objective performance indicators and subjective ratings. Pilots flew a high-workload manual control mission and a high-workload mission that emphasized mental activities. Subjective ratings for the two missions were different, but the objective performance measures (altitude deviations) were similar. In the second set of experiments, workload levels were increased and a second performance measure was taken (e.g., airspeed deviations). Mental workload had no influence on either performance-based workload measure. Subjective ratings discriminated among the scenarios and correlated with performance measures for high-workload flights. The number of mental tasks performed did not influence error rates, although high manual workloads did increase errors.

015 Berg, S. L., & Sheridan, T. B. (1986). The impact of physical and mental tasks on pilot mental workload. Proceedings of the 21st Annual Conference on Manual Control, NASA CP-2428, pp. 6.1-6.26.

Seven instrument-rated pilots with a wide range of backgrounds and experience levels flew four different scenarios on a fixed-base simulator. The Baseline scenario was the simplest of the four and had few mental and physical tasks. An Activity scenario had many physical but few mental tasks. The Planning scenario had few physical and many mental tasks. A Combined scenario had high mental and physical task loads. The magnitude of each pilot's altitude and airspeed deviations was measured, subjective workload ratings were recorded, and the degree of pilot compliance with assigned memory/planning tasks was noted. Mental and physical performance was a strong function of the manual activity level, but not influenced by the mental task load. High manual task loads resulted in a large percentage of mental errors even under low mental task loads. Although all the pilots gave similar subjective ratings when the manual task load was high, subjective ratings showed greater individual differences with high mental task loads. Altitude or airspeed deviations and subjective ratings were most correlated when the total task load was very high. Although airspeed deviations, altitude deviations, and subjective workload ratings were similar for both low experience and high experience pilots, at very high total task loads, mental performance was much lower for the low experience pilots.

016 Biferno, M. A. (1985). Mental workload measurement: Event-related potentials and ratings of workload and fatigue (NASA CR-177354). Washington, DC: National Aeronautics and Space Administration.

Event-related potentials were elicited when a digitized work representing a pilot's call-sign was presented. This auditory probe was presented during 27 workload conditions in a 3 x 3 x 3 design where the following variables were manipulated: short-term memory load, tracking task difficulty, and time-on-task. Ratings of workload and fatigue were obtained between each trial of a 2.5-hour test. The data of each subject were analyzed individually to determine whether significant correlations existed between subjective ratings and ERP component measures. Results indicated that a significant number of subjects had positive correlations between: (1) ratings of workload and P300 amplitude, (2) ratings of workload and N400 amplitude, and (3) ratings of fatigue and P300 amplitude. These data are the first to show correlations between ratings of workload or fatigue and ERP components thereby reinforcing their validity as measures of

mental workload and fatigue. Since ratings of fatigue and workload were significantly correlated for 16 of 20 subjects, future studies of workload would benefit from examining the relationship between them.

017 Bloem, K. A., & Damos, D. L. (1985). Individual differences in secondary task performance and subjective estimation of workload. Psychological Reports, 56, 311-322.

This experiment had two purposes. First, it attempted to replicate the easy-to-hard prediction for residual capacity described by Lansman and Hunt (1982) for two complex task combinations. Second, it examined the relation between individual differences in resource capacity, as indicated by the easy-to-hard prediction, and the subjective experience of workload. One task combination involved a verbal memory task paired with a vowel-consonant classification task. The other combination involved a paired associate task with a name classification task. The easy-to-hard prediction was not replicated for either task combination; easy primary task performance provided a better prediction of hard primary task performance than did secondary task performance. Measures of residual capacity were not related to subjective ratings of workload, however, the workload scales were sensitive to between-task differences.

(Reprinted by permission.)

018 Bortolussi, M. R., Kantowitz, B. H., & Hart, S. G. (1985). Measuring pilot workload in a motion base trainer: A comparison of four techniques. Proceedings of the Third Symposium on Aviation Psychology, pp. 263-270.

Various techniques have been developed to predict and measure pilot workload. This simulation was conducted in order to compare four widely used methods: A visual two- and four-choice reaction time task, time production, retrospective multi-dimensional subjective ratings and in-flight verbal workload estimates. Two scenarios with different levels of difficulty determined by preliminary research were designed to test these techniques. The insertion of the secondary tasks did not significantly affect flight performance. All four techniques were able to distinguish among levels of scenario complexity. In addition, the three secondary tasks and workload ratings obtained in-flight were generally able to distinguish among levels of difficulty for different segments within the scenarios.

019 Boyd, S. P. (1983). Assessing the validity of SWAT as a workload measurement instrument. Proceedings of the Human Factors Society 27th Annual Meeting, 1, 124-128.

The Subjective Workload Assessment Technique (SWAT) carries with it the implicit assumption that people can accurately predict the amount of mental workload they would experience under various levels of three component dimensions. Research suggests that the perceptions of these dimensions may not be independent. This study was designed to measure the subjective interactions between the dimensions used in SWAT. Mental workload was generated using a text editing task in which the dimensions were manipulated independently. Results revealed significant positive correlations between the subjective levels of the three dimensions.

That is, when a subject experienced a high level of one dimension, s/he also tended to rate the other two dimensions high. It may be unreasonable to assume that people can accurately predict the magnitude of these interactions when performing the ranking process which is used to derive the workload scale.

(From Proceedings of the Human Factors Society 27th Annual Meeting, 1983, 1, p. 124. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

020 Braune, R., & Wickens, C. D. (1983). Individual differences and age-related changes in time-sharing ability of aviators. Proceedings of the Human Factors Society 27th Annual Meeting, 1, 117-121.

Performance in single task and dual task configurations was examined across four different age groups to determine the presence of an attention deficit hypothesis with increasing chronological age. Although a general slowing of performance could be shown, no interaction between age and dual task loadings could be observed which is interpreted as negative evidence for the attention deficit hypothesis. A separate analysis revealed individual differences in time-sharing ability within age groups to be a significant factor in dual task performance.

(From Proceedings of the Human Factors Society 27th Annual Meeting, 1983, 1, p. 117. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

021 Burford, C. L., Ramsey, J. D., Tayyari, F., Lee, C. H., & Stepp, R. G. (1984). A method for systematic workload estimation (SWE). Proceedings of the Human Factors Society 28th Annual Meeting, 2, 997-999.

Using the information available in the literature, a coding method has been developed which reduces the difficulties associated with the estimation of metabolic workloads. A person can be trained to use the system with minimum time and effort. The only equipment required are worksheets, watch and a clip board. Systematic Workload Estimation (SWE) can be applied to the activities with constant or changing paces. Its validity has been examined by coding and calculating the metabolic costs of numerous activities and comparing the results with the values given in published tables.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 2, p. 997. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

022 Butterbaugh, L. C., & McBride, J. K. (1984). A computer-based system for workstation design and analysis. Proceedings of the Human Factors Society 28th Annual Meeting, 1, 63-66.

The Crew Systems Development Branch of the US Air Force Wright Aeronautical Laboratories is pursuing a system of computer aided design and evaluation techniques (CADETs) for aircraft crew station development. Composed of individual computer based human factors tools and anaylsis methods, the system's key element is a user interface module (UIM). The UIM serves as the single user-

computer interface through which all CADETs are accessed, eliminating many of the potential disadvantages normally associated with such an approach (e.g. diverse user backgrounds and varying user familiarity with computers). The CADETs are being implemented in a modular architecture to permit easy upgrade, replacement, or addition of any design tool or analysis techniue. These features, and others discussed herein, are being combined to produce a system with numerous desirable attributes including user friendliness, and software portability. A powerful, effective CADET system will result, employing computer graphic, analytic, and data base technologies for the design of workstation layouts, and the analysis of workstation geometry, man-machine function allocation, and operator workload.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 1, p. 63. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

023 Butterbaugh, L., Warner, D., Lovering, P., & Herron, S. (1981). Pilot workload: A survey of operational problems (AFWAL-TR-81-3011). Wright-Patterson Air Force Base, OH: Air Force Wright Aeronautical Laboratories. (DTIC No. AD-A107 758)

Five hundred and seventy three USAF pilots responded to a survey which had as its objective the identification of operational, crew station design related causes of high pilot workloads. The survey consisted of mailed survey forms and personal interviews structured to conform with the "critical incident technique" of collecting user-provided data. The survey canvased over 50 USAF organizations in collecting data for more than 30 currently flown USAF aircraft types. The role control/display design, crew station design, and equipment malfunctions play in contributing to cockpit workload is unique to each aircraft. Other factors, such as weather, flight schedules, and mission phase appear to contribute to cockpit workloads in most all the aircraft surveyed. Further, the reported situations, or "critical incidents", indicate that high workloads result from the simultaneous occurrence or existence of several causes. For example, a high workload situation reported for the FB-111 consisted of an equipment failure while flying the low-level penetration portion of a mission at night. All data collected has been catalogued for the establishment of an information base, and available for future use in conjunction with aircraft development programs or modernization/retrofit programs.

024 Casali, J. G., & Wierwille, W. W. (1982). A sensitivity/intrusion comparison of mental workload estimation techniques using a flight task emphasizing perceptual piloting activities. Proceedings of the 1982 IEEE International Conference on Cybernetics and Society, pp. 598-602.

There are many flight task situations in which perceptual activity on the part of the pilot or aircrew member is emphasized. Unfortunately, the sensitivity, that is, the relative ability of conventional workload estimation techniques to discriminate between perceptual load levels, is largely unknown. Because of this lack of basic knowledge, an experiment comparing several workload techniques was

conducted in an instrumented GAT-1B flight simulator. The initial sensitivity and intrusion results of the experiment are reported in this paper, and a relative categorization of techniques is presented, based on demonstrated sensitivity.

(Copyright 1982 IEEE. Reprinted by permission from Proceedings of the 1982 IEEE International Conference on Cybernetics and Society, Seattle, WA, October 28-30, p. 598.)

025 Casali, J. G., & Wierwille, W. W. (1983). A comparison of rating scale, secondary-task, physiological, and primary-task workload estimation techniques in a simulated flight task emphasizing communications load. Human Factors, 25, 623-641.

Sixteen potential metrics of pilot mental workload were investigated regarding their sensitivity to communication load and their intrusion on primary-task performance. A moving-base flight simulator was used to present three cross-country flights. The flights varied only in the difficulty of the communications requirements. Rating scale measures were obtained immediately postflight; all others were taken over a 7-min segment of the flight task. The results indicated that both the Modified Cooper-Harper scale and the workload Multi-descriptor Scale were sensitive to changes in communications load. The secondary-task measure of time estimation and the physiological measure of pupil diameter were also sensitive. As expected, those primary-task measures that were direct measures of communicative performance were also sensitive to load, whereas aircraft control primary-task measures were not, attesting to the task specificity of such measures. Finally, the intrusion analysis revealed no differential interference between workload measures.

(From Human Factors, 1983, 25, p. 623. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

026 Casali, J. G., & Wierwille, W. W. (1984). On the measurement of pilot perceptual workload: A comparison of assessment techniques addressing sensitivity and intrusion issues. Ergonomics, 27, 1033-1050.

A flight simulator-based study was conducted to examine fourteen distinct mental workload estimation measures, including opinion, secondary task, physiological, and primary task measures. Both the relative sensitivity of the measures to changes in mental workload and the differential intrusion of the changes on primary task performance were assessed. The flight task was varied in difficulty by manipulation of the presentation rate and complexity of a hazard-perception task that required each of 48 licensed pilots to rely heavily on their perceptual abilities. Three rating scales (Modified Cooper-Harper, Multidescriptor, and Workload-Compensation-Interference/Technical Effectiveness), two secondary task measures (time estimation and tapping regularity), one physiological measure (respiration frequency), and one primary task measure (danger-condition response time) were reliable indicants of workload changes. Recommendations for applying the workload measures are presented.

(Reprinted by permission.)

027 Casper, P. A., Shively, R. J., & Hart, S. G. (1986). Workload consultant: A microprocessor-based system for selecting workload assessmen procedures. Proceedings of the 1986 IEEE International Conference on System , Man, and Cybernetics, 2, 1054-1059.

Recent years have seen a deepening interest in the measurement of human operator workload. However, not all persons involved in the design and production of human-machine systems are educated in the rigors of workload measurement and the currently available techniques. Furthermore, as in most areas of expertise, there aren't enough human "experts" to go around. The present paper describes an "expert" system, created at the NASA-Ames Research Center, that was designed to provide decision support for persons interested in assessing operator workload. The system is based on current research in the field of workload measurement and is flexible enough to allow for incorporation of new knowledge as it is empirically validated.

028 Charlton, S. G. (1985). Behavior analysis: A tool for test and evaluation. Proceedings of the Human Factors Society 29th Annual Meeting, 1, 188-192.

The following paper describes an approach for measuring the behavioral consequences of a wide variety of factors in a field test environment. Behavior analysis provides a test and evaluation team with an indication of what behavior changes occur concomitant with high workload, stress, and fatigue. As such, this assessment technique is used in conjunction with other evaluation tools to enable evaluation of any divergence in the behavior of the personnel manning the system (i.e., deviations from either specified or previous behavioral performance.).

029 Chechile, R. A., & Sadoski, D. M. (1983). The effects of cuing in time-shared tasks. Human Factors, 25, 371-377.

Two divided-attention experiments were performed to evaluate the effectiveness of verbal and symbolic cuing on the primary task of editing flight route-way-point information. Cuing conditions were contrasted with a no-cue control. Symbolic cuing consisted of directing the operators to the appropriate keys throughout the sequence of primary task steps. Verbal cuing consisted of a short instruction displayed on the CRT just prior to the next step. In Experiment 1, symbolic cuing proved effective in improving editing accuracy. In Experiment 2, symbolic cuing resulted in reduced early-trial edit-completion time. Verbal cuing,

however, was not effective on these tasks in either experiment. It is theorized that symbolic cuing lightened the cognitive load, whereas verbal cuing did not reduce the cognitive load since it still required attention to read the verbal message.

(From Human Factors, 1983, 25, p. 371. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

030 Childress, M. E. (1983). Subjective scales for workload evaluation: Critical aspects and new directions for research. Proceedings of the 19th Annual Conference on Manual Control, pp. 1-2.

As aircraft and other mechanical systems increase in complexity and rely more heavily on computerization of function, and as the pilot or other operator assumes greater supervisory responsibility for system monitoring and control, need for evaluation of the workload associated with system changes increases. Many of the methods currently available, though helpful in specific situations and often necessary in promoting understanding of some basic processes, are often difficult and unwieldy to use in complex, practical situations. Subjective rating scales, however, are convenient instruments for evaluating this workload and for estimating the magnitude of changes in load as system changes occur. The use of such scales has historical precedent in the personnel literature, particularly in performance evaluation. Subjective scales also have been used to evaluate specific system characteristics, such as aircraft handling qualities. The utility of the method is clear; however, psychometric development of subjective scales for the evaluation of workload currently is in its infancy. Thus though the literature is replete with examples of and recommendations for their use as well as with criticisms of their deficiencies, research directed toward examination of their properties, and evaluation of the conditions under which their use is appropriate and obtained results generalizable is just beginning. Several important works (e.g., Landy and Farr, 1980; Moray, 1979; Nisbett and Wilson, 1977; Wherry, 1950, 1952) have described the problems associated with subjective ratings, have detailed some of the situations in which they may be appropriate, and have recommended specific topics for future research. This paper presents a review of critical aspects of that literature which suggest directions for future research relative to self-ratings of subjective workload. It provides examples of some recent work at Ames Research Center which has suggested extending the basic input-processing-outcome model for examining workload to consider all input sources and the related outcomes, and it details current work based on that model.

031 Childress, M. E., Hart, S. G., & Bortolussi, M. R. (1982). The reliability and validity of flight task workload ratings. Proceedings of the Human Factors Society 26th Annual Meeting, pp. 319-323.

Twelve instrument-rated general aviation pilots each flew two scenarios in a motion-base simulator. During each flight, the pilots verbally estimated their workload every three minutes. Following each flight, they again estimated workload for each flight segment and also rated their overall workload, perceived

performance, and 13 specific factors on a bipolar scale. The results indicate that time (a priori, inflight, or postflight) of eliciting ratings, period to be covered by the ratings (a specific moment in time or a longer period), type of rating scale, and rating method (verbal, written, or other) may be important variables. Overall workload ratings appear to be predicted by different specific scales depending upon the situation, with activity level the best predictor. Perceived performance seems to bear little relationship to observer-rated performance when pilots rate their overall performance and an observer rates specific behaviors. Perceived workload and performance also seem unrelated.

(From Proceedings of the Human Factors Society 26th Annual Meeting, 1982, p. 319. Copyright 1982 by the Human Factors Society, Inc. Reprinted by permission.)

032 Chiles, W. D., Ellis, G. A., & Roscoe, A. H. (1978). Assessing pilot workload (AGARD-AG-233). Neuilly-Sur-Seine, France: North Atlantic Treaty Organization - Advisory Group for Aerospace Research and Development. (DTIC No. AD-A051 587)

The assessment of levels of pilot workload associated with the various phases and sub-phases of flight is important in the design, development, and evaluation of aircraft handling qualities and of display and guidance systems. This AGARDograph, written primarily for flight test engineers and pilots, is intended as a guide to the different methods available for estimating workload and in particular to those techniques suitable for use in aircraft. An introductory chapter briefly reviews the various concepts and classifications of workload; the former tend to fall into two main areas, those related to workload as task-demands and those to workload as pilot-effort. In Chapter 2, subjective assessment, at present the most used method, is discussed from the viewpoint of the test pilot. Physiological methods in general are reviewed in Chapter 3 with those techniques available for use in flight being discussed in more detail. Chapter 4 describes various objective methods and presents examples of their practical application. Whereas the methods in Chapter 2 and 3 are appropriate only to workload as effort, objective methods contain techniques appropriate to workload as task-demands as well as to effort. The former techniques are particularly valuable for providing data which can be used to construct models and to predict levels of workload. Different modelling techniques will be discussed in a proposed supplement entitled Engineering Methods.

(The original version of this material was first published by the Advisory Group for Aerospace Research and Development, North Atlantic Treaty Organization [AGARD/NATO] in AG 233, Assessing Pilot Workload, February 1978. Reprinted by permission.)

033 Ciavarelli, A. P., & Brictson, C. A. (1978). Air combat maneuvering range (ACMR): Has operational performance measurement entered a golden age? Proceedings of the Human Factors Society 22nd Annual Meeting, pp. 365-368.

Three years of aircrew performance measurement related to air combat effectiveness using the Navy's Air Combat Maneuvering Range (ACMR) are described. Performance assessment methods were based on air combat engagement outcomes

(i.e., wins, losses, draws), weapon delivery accuracy measures, and metrics derived from antecedent events. When used in an operational setting, the aircrew assessment methods have been used to identify squadron performance differences, evaluate competitive exercises, and provide diagnostic training feedback to operational users. The use of continuously recorded quantitative measures from systems such as ACMR represents a 'Golden Age' in the performance measurement field. The availability of objective performance criteria promises to be of substantial benefit to both the operational user and the research community in such areas as pilot selection and training, fleet combat readiness and pilot workload and stress.

(From Proceedings of the Human Factors Society 22nd Annual Meeting, 1978, p. 365. Copyright 1978 by the Human Factors Society, Inc. Reprinted by permission.)

034 Cohen, A. D. (1982). The Hughes Design Analysis System and instructor workload in operational trainers. Proceedings of the Human Factors Society 26th Annual Meeting, pp. 364-368.

Computerized simulator trainers require varying degrees of instructor participation in training exercises and student performance evaluation. An estimation of instructor workload before the system has been fully developed is essential for an effective functional allocation of human and computer capabilities. One approach for estimating operator workload is simulation. The Hughes Design Analysis System (DAS) is an interactive and graphic simulation package. The present paper describes: (1) how DAS has been used for an instructor workload analysis, (2) the analysis method, (3) the results, and (4) the recommendations given to system designers.

(From Proceedings of the Human Factors Society 26th Annual Meeting, 1982, p. 364. Copyright 1982 by the Human Factors Society, Inc. Reprinted by permission.)

035 Colle, H. A. (1979). A capacity-theoretic approach to workload assessment (WSU-AFOSR-TR-79-1). Dayton, OH: Wright State University. (DTIC No. AD-A078 132)

Secondary task measurement of mental workload was explored both theoretically and empirically. By analyzing sets of secondary tasks over broad performance ranges, it was found that an additive scale of workload was feasible. Therefore, rational and empirical bases exist for using secondary tasks to measure the workload during flying tasks or other complex skills. A formal measurement theory for testing additivity and a scheduling theory of workload were described.

(Reprinted by permission.)

036 Connor, S. A., & Wierwille, W. W. (1983). Comparative evaluation of twenty pilot workload assessment measures using a psychomotor task in a moving-base aircraft simulator (NASA CR-166457). Washington, DC: National Aeronautics and Space Administration.

A comparison of the sensitivity and intrusion of twenty pilot workload assessment techniques was conducted using a psychomotor loading task in a three degree of freedom moving-base aircraft simulator. The twenty techniques included opinion measures, and primary task performance measures. The primary task was an instrument landing system (ILS) approach and landing. All measures were recorded between the outer marker and the middle marker on the approach. Three levels (low, medium, and high) of psychomotor load were obtained by the combined manipulation of windgust disturbance level and simulated aircraft pitch stability. Six instrument rated pilots participated in four sessions lasting approximately three hours each.

037 Cote, D. O., Krueger, G. P., & Simmons, R. R. (1985). Helicopter copilot workload during nap-of-the-earth flight. Aviation, Space, & Environmental Medicine, 56, 153-157.

Two automatic navigation systems, a Doppler radar system and a projected map system, and a hand-held map were examined for their effects on copilot/navigator workload and performance. The automatic navigation systems reduced the number of navigation errors and the size of deviations from intended track. The Doppler system reduced the time devoted to navigating and the number of verbal navigation messages exchanged between the pilot and copilot. The projected map system reduced visual workload. However, with all three navigation systems, more than 80% of the copilots' time was spent on navigation tasks, less than 10% of their time was visual "free time" that could be used for other tasks, and greater than 20% of the aircrew's time was occupied with navigation communications.

(Reprinted by permission.)

038 Courtright, J. F., & Kuperman, G. (1984). Use of SWAT in USAF system T&E. Proceedings of the Human Factors Society 28th Annual Meeting, 2, 700-703.

Many of today's complex military systems place critical performance demands upon the men and women who must operate them. Operational test and evaluation of these systems has long required the assessment of operator behavior within the context of system performance. Interest in estimating the effect on operator performance of system demands not specifically examined during the test led to adding assessment of operator workload to the T&E process. The Subjective Workload Evaluation Technique (SWAT) rating scale was employed in a field evaluation of a system requiring skilled personnel to operate semi-automated equipment. Discussed are the means employed to identify tasks of interest, the acceptance of SWAT by the operators performing the ratings, and the nature of the system problems identified by use of the technique.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 2, p. 700. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

039 Crabtree, M. S., Bateman, R. P., & Acton, W. H. (1984). Benefits of using objective and subjective workload measures. Proceedings of the Human Factors Society 28th Annual Meeting, 2, 950-953.

This paper describes the results of two measures of workload, one objective, one subjective, that were applied to a series of switch setting exercises. The objective measure was an Interval Production Task (IPT), based upon a secondary task paradigm, in which the subject is required to maintain a constant rate of tapping with one hand while performing the primary task with the other hand. According to the supporting theory, variations in workload on the primary task will be reflected in variations in the tapping rate. Previous studies have found the IPT to be particularly sensitive to changes in the psychomotor workload. The subjective measures used were the Subjective Workload Assessment Technique (SWAT). Although the two workload measures did not produce the same level of significance for workload differences, it should be noted that they did produce the same rank order for the three tasks. It was concluded that the use of both subjective and objective measures of workload could produce increased confidence in the results, as well as insight into the nature of the task loading.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 2, p. 950. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

040 Crombie, R. B. (1982). Reflections on the effects of vehicle dynamics and task difficulty on Cooper-Harper pilot opinion ratings, task performance, and pilot workload. In M. L. Frazier & R. B. Crombie (Eds.), Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics (AFFTC-TR-82-5, pp. 102-113). Edwards Air Force Base, CA: Air Force Flight Test Center. (DTIC No. AD-A129 333)

This paper is intended to set forth generalized principles on a broad subject which lies at the heart of both flying qualities and human factors research, development, test, and evaluation. These generalizations are the result of the author's study of the flying qualities, pilot workload, and pilot performance literature. They are submitted to the participants of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics to stimulate further discourse on the application of these principles to particular experimental circumstances and to encourage the readers to recognize the common threads linking the fields of aircraft flying qualities and human factors.

041 Damos, D. (1984). Classification schemes for individual differences in multiple-task performance and subjective estimates of workload. Proceedings of the 20th Annual Conference on Manual Control, NASA CP-2341, 2, 97-104.

Human factors practitioners often are concerned with mental workload in multiple-task situations. Investigations of these situations have demonstrated repeatedly that individuals differ in their subjective estimates of workload. These differences may be attributed in part to individual differences in definitions of workload (Hart, Childress, and Hauser, 1982). However, after

allowing for differences in the definition of workload, there are still unexplained individual differences in workload ratings. The general purpose of the two studies reported in this paper was to examine the relation between individual differences in multiple-task performance, subjective estimates of workload, information processing abilities, and the Type A personality trait.

042 Damos, D. (1984). Subjective workload and individual differences in information processing abilities. Proceedings of the Fourth Annual Aerospace Behavioral Engineering Technology Conference, pp. 71-74.

This paper describes several experiments examining the source of individual difference in the experience of mental workload. Three sources of such differences were examined: information processing abilities, timesharing abilities, and personality traits/behavior patterns. On the whole, there was little evidence that individual differences in information processing abilities or timesharing abilities are related to perceived differences in mental workload. However, individuals with strong Type A coronary prone behavior patterns differed in both single- and multiple-task performance from individuals who showed little evidence of such a pattern. Additionally, individuals with a strong Type A pattern showed some dissociation between objective performance and the experience of mental workload.

043 Damos, D., & Smist, T. (1981). Individual differences in multi-task response strategies. Proceedings of the Human Factors Society 25th Annual Meeting, pp. 291-295.

Two experiments are presented demonstrating that the response strategies used to perform a discrete task combination reflect significant individual differences in multiple-, but not single-, task performance and that these differences are evident in a variety of task combinations.

044 Damos, D. L. (1984). Examining the relation between subjective estimates of workload and individual differences in performance (NASA CR-2341). Washington, DC: National Aeronautics and Space Administration.

The primary purpose of this 2-year grant was to examine the relation between subjective estimates of workload, personality measures, and individual differences in single- and multiple-task performance. As specified in the grant proposal and second-year revision, four experiments were completed during the course of the grant examining these relations.

045 Damos, D. L. (1984). Individual differences in multiple-task performance and subjective estimates of workload. Perceptual and Motor Skills, 59, 567-580.

This experiment examined the relation between individual differences in multiple-task performance and subjective estimates of workload. 30 women performed various complex tasks alone and together and rated each task and task combination of 10 bipolar adjective scales describing different dimensions of workload. The subjects also completed tests of field dependence, memory span, and time estimation. Two classification schemes were used to identify each subject. One was based on the subject's dual-task response strategy and the other, on the subject's performance on a complex monitoring task. As the data showed little evidence of consistent individual differences on the monitoring task, this classification system was subsequently dropped. Group differences in between-response strategies were found on two of the workload scales. Some between-group trends were found on the time estimation and memory-span tasks, suggesting additional topics for investigation.

(Reprinted by permission.)

046 Damos, D. L. (1985). The relation between the type A behavior pattern, pacing, and subjective workload under single and dual-task conditions. Human Factors, 27, 675-680.

Twenty Type A and 20 Type B subjects performed two discrete tasks alone and together. Half of the subjects performed paced versions of both tasks; half, unpaced versions. Workload ratings were obtained for all subjects under single- and dual-task conditions using eight bipolar adjective scales. Under single-task conditions there was a significant interaction between behavior pattern and pacing on one of the tasks. This interaction indicated that Type A subjects responded more rapidly under unpaced conditions than did Type B subjects, although there was little difference between the groups under paced conditions. Under dual-task conditions, Type A subjects responded more rapidly than did Type B subjects regardless of pacing. There was one significant interaction between behavior pattern and task on one of the workload scales.

(From Human Factors, 1985, 27, p. 675. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

047 Damos, D. L., & Lintern, G. (1980). A comparison of the predictive validities of single- and dual-task measures. Proceedings of the Human Factors Society 24th Annual Meeting, pp. 245-248.

An experiment comparing the predictive validity of single- versus dual-task measures is reported. Fifty-seven males received two trials on each of two identical one-dimensional compensatory tracking tasks followed by 25 dual-task trials. Finally, they performed each task alone for one trial. The subjects then were given a short basic flight course consisting of ground instruction and practice in a GAT-2 simulator. After completing the course, the subjects performed four repetitions of three maneuvers. Performance in the simulator then was correlated with performance on each tracking trial. The predictive validity

of the early single-task scores decreased with practice while the dual-task validity increased throughout the testing session. However, the predictive validity of the late single-task scores was almost as large as that of the late dual-task scores.

(From Proceedings of the Human Factors Society 24th Annual Meeting, 1980, p. 245. Copyright 1980 by the Human Factors Society, Inc. Reprinted by permission.)

048 De Maio, J., Bell, H. H., & Brunderman, J. (1983). Pilot oriented performance measurement. Proceedings of the Human Factors Society 27th Annual Meeting, 1, 463-467.

Flight simulators provide a complete quantitative record of a pilot's flying performance. Evaluating this record is complicated by the volume of data and by its fine detail, dozens of flight parameters, sampled many times per second. Automated performance measurement systems (APMS) reduce the volume of data to an amount which is manageable and understandable. The usual APMS is aircraft state oriented. The APMS keys on aircraft state (e.g., X-Y position, bank angle) to define intervals over which performance data are integrated. This APMS is relatively insensitive to pilots' intentions and so may average performances which had differing objectives, based only on their having occurred at the same point during the task sequence. An alternative APMS has been developed which is pilot oriented. This APMS defines measurement intervals based on control inputs. Control inputs are identified by discrete changes in flight path. These intervals are psychologically relevant in that they begin with a goal- directed control input and end with a countervailing input. By relating performance in the pilot defined intervals to state defined intervals, it is posible to quantify performance on given flight segments (e.g., a level turn), and to specify factors which lead to a given level of performance.

(From Proceedings of the Human Factors Society 27th Annual Meeting, 1983, 1, p. 463. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

049 Derrick, W. L. (1981). The relationship between processing resource and subjective dimensions of operator workload. Proceedings of the Human Factors Society 25th Annual Meeting, pp. 532-536.

A multiple structure model of processing resources (Wickens, 1980) guided construction of tasks of differential resource demand that were both performed by subjects and rated according to workload similarity. Analysis of performance data generally supported model predictions. Multidimensional scaling analysis of the similarity data produced subjective dimensions of workload that were explained in terms of resource demand and task structure.

(From Proceedings of the Human Factors Society 25th Annual Meeting, 1981, p. 532. Copyright 1981 by the Human Factors Society, Inc. Reprinted by permission.)

050 Derrick, W. L. (1983). Examination of workload measures with subjective task clusters. Proceedings of the Human Factors Society 27th Annual Meeting, 1, 134-138.

Multiple measures of operator workload may fail to agree or dissociate for a given task. This study proposes a new method to examine this dissociation for two categories of workload measures: Subjective ratings and performance-based secondary tasks. Eighteen tasks of differential processing resource demand were performed by subjects and rated according to workload similarity. Additive clustering analysis of the workload ratings produced overlapping task clusters. Three properties--performance, effort, and input complexity--explained the cluster solution. Dissociation was found when tasks perceived as similar in workload did not possess the same properties.

(From Proceedings of the Human Factors Society 27th Annual Meeting, 1983, 1, p. 134. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

051 Derrick, W. L., & McCloy, T. M. (1984). An empirical demonstration of multiple resources. Proceedings of the Human Factors Society 28th Annual Meeting, 1, 26-30.

Earlier work by Gopher, Brickner, and Navon (1982) suggested that evidence for multiple processing resources can be demonstrated only if both task difficulty and task pricrities are manipulated in dual task studies. To further investigate their approach, ten subjects performed both single and dual task versions of a tracking task and a vowel insertion task, the latter modified to increase either motor load or cognitive load. Dual task trials required equal task emphasis on the favoring of one task over another. As expected, a difficulty by priority interaction was found in the motor load but not in the cognitive load condition, suggesting multiple resources. Performance Operating Characteristics for these data suggested that more than two resources were being utilized by these tasks.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 1, p. [illegible] Copyright 1984 by the Human Factors Society, Inc. Reprinted by [illegible]on.)

052 [illegible]unn, R. S. (1982). Army workload research and development requirements. In M. L. Frazier & R. B. Crombie (Eds.), Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics (AFFTC-TR-82-5, pp. 20-25). Edwards Air Force Base, CA: Air Force Flight Test Center. (DTIC No. AD-A129 333)

This article presents a discussion of the principle current Army requirement for research and methodology development in pilot and aircrew workload measurement. Emphasis is given to the shifting impact of operator task demands and to the utility value of workload methodology in guiding design and development activities. The discussion focuses on requirements which are especially acute in the system acquisition process involving development of new helicopter crew stations. Conclusions and opinions offered by the writer suggest both a conceptual orientation and a general approach to meeting the requirements.

053 Eckel, J. S., & Crabtree, M. S. (1983). Analytic and subjective assessments of operator workload imposed by communications tasks in transport aircraft. Proceedings of the Second Symposium on Aviation Psychology, pp. 237-241.

The purpose of this project is to use analytical and subjective techniques to estimate the workload imposed on the aircrew by typical communications-related tasks performed during selected flight phases. Communications-related tasks are defined operationally to consist of sequences of verbal and discrete manual responses which are initiated when the crew receives and interprets radio messages from an Air Traffic Control (ATC) Facility. Three evaluation techniques will be used to quantify communications-related workload. The first, an information theoretic technique, permits determination of bit values for perceptual and for verbal and manual action components of each task. The second is a paired comparison technique to obtain subjective estimates of the cognitive processing demands of individual communication requests. By combining the results of the paired comparison analysis with the results of the information theoretic analysis, we will derive a single hybrid scale of communications-related workload. The third technique relies on pilots' estimations of the overall workload associated with communications tasks. Recommendations for future research include an examination of communications workload among the air crew and the development of simulation scenarios which impose distinctly different levels of communications-related workload.

054 Eggemeier, F. T. (1980). Some current issues in workload assessment. Proceedings of the Human Factors Society 24th Annual Meeting, pp. 669-673.

Two current issues related to operator workload assessment include the effects of operator strategy on levels of load experienced by the operator and the nature of operator information processing capacities and resources. Several current positions related to each of these issues are presented, and implications for workload assessment are discussed. It is concluded that workload is a multidimensional construct which requires the development of an assessment battery which should include subjective, performance-based, and physiological measures.

(From Proceedings of the Human Factors Society 24th Annual Meeting, 1980, p. 669. Copyright 1980 by the Human Factors Society, Inc. Reprinted by permission.)

055 Eggemeier, F. T. (1981). Current issues in subjective assessment of workload. Proceedings of the Human Factors Society 25th Annual Meeting, pp. 513-517.

Because of their practical utility, subjective measures have been widely used as workload assessment techniques. There has also been increased support in the recent literature for the importance of including subjective measures as part of a comprehensive workload assessment methodology. In spite of the widespread use and importance of subjective techniques, several significant issues which pertain to their development and use remain unresolved. One such issue is related to the

development of a generalized subjective workload assessment procedure, while a second issue concerns the need to identify factors or variables that contribute to the subjective experience of workload. This paper discusses both issues and outlines some recent research which is applicable to each area.

(From Proceedings of the Human Factors Society 25th Annual Meeting, 1981, p. 513. Copyright 1981 by the Human Factors Society, Inc. Reprinted by permission.)

056 Eggemeier, F. T., Crabtree, M. S., & LaPointe, P. A. (1983). The effect of delayed report on subjective ratings of mental workload. Proceedings of the Human Factors Society 27th Annual Meeting, 1, 139-143.

Forty-eight subjects performed a short-term memory task with several difficulty levels and provided either immediate or delayed ratings of workload via the Subjective Workload Assessment Technique (SWAT). Mean SWAT ratings did not vary significantly as a function of delayed report, but a substantial number of subjects gave delayed ratings that were discrepant from their immediate ratings. A counterbalancing effect in delayed ratings appears to have been a factor in the failure of the delay effect to reach significance. A secondary objective of this study was to examine the sensitivity of SWAT in a between-subjects design. SWAT ratings varied significantly as a function of task difficulty manipulations, supporting the sensitivity of SWAT to the workload of the conditions used.

(From Proceedings of the Human Factors Society 27th Annual Meeting, 1983, 1, p. 139. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

057 Eggemeier, F. T., Crabtree, M. S., Zingg, J. J., Reid, G. B., & Shingledecker, C. A. (1982). Subjective workload assessment in a memory update task. Proceedings of the Human Factors Society 26th Annual Meeting, pp. 643-647.

Twelve subjects performed a short-term memory task under several difficulty levels and rated the workload associated with each condition using the Subjective Workload Assessment Technique (SWAT). SWAT ratings proved more sensitive than memory error to task difficulty variations in all but one of the most difficult memory conditions. Most importantly, SWAT ratings demonstrated their greatest relative sensitivity at the lowest levels of workload. The results are interpreted as supporting the applicability of SWAT as a sensitive workload index.

(From Proceedings of the Human Factors Society 26th Annual Meeting, 1982, p. 643. Copyright 1982 by the Human Factors Society, Inc. Reprinted by permission.)

058 Eggemeier, F. T., Melville, B. E., & Crabtree, M. S. (1984). The effect of intervening task performance on subjective workload ratings. Proceedings of the Human Factors Society 28th Annual Meeting, 2, 954-958.

Thirty subjects performed a short-term memory task and used the Subjective Workload Assessment Technique (SWAT) to provide workload ratings under one of five conditions. Ratings were provided either immediately following task performance, after a delay period during which no additional tasks were performed, or after a delay period during which an additional set of memory tasks at one of three levels of difficulty was performed. Neither the delay interval nor the requirement to perform a set of intervening tasks significantly affected mean SWAT ratings relative to the immediate rating control condition. Patterns in the data suggested that performance of a set of difficult intervening tasks had the greatest tendency to affect memory task ratings, and indicate that the potential influence of intervening task performance should not be completely discounted in workload rating scale applications.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 2, p. 954. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

059 Eggemeier, F. T., Shingledecker, C. A., & Crabtree, M. S. (1985). Workload measurement in system design and evaluation. Proceedings of the Human Factors Society 29th Annual Meeting, 1, 215-219.

Because of its central role in system development, workload measurement has been extensively researched. These efforts have produced a variety of workload assessment techniques, many of which can be classified as either subjective, physiological, or behavioral measures. These categories of measures can vary along several dimensions that can be used as criteria in selection of a technique for a particular application. The proposed selection criteria include the sensitivity, diagnosticity, and intrusiveness associated with a technique. Different stages of system design can require techniques that differ on the noted dimensions. Since no technique is capable of meeting all of the applicable criteria, a comprehensive approach to workload assessment will require a battery of subjective, physiological, and behavioral measures. Future research dealing with comparative evaluation of the various assessment techniques along the noted dimensions will be required in order to refine workload metric selection criteria.

(From Proceedings of the Human Factors Society 29th Annual Meeting, 1985, 1, p. 215. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

060 Eggemeier, F. T., & Stadler, M. A. (1984). Subjective workload assessment in a spatial memory task. Proceedings of the Human Factors Society 28th Annual Meeting, 2, 680-684.

Twelve subjects performed a spatial short-term memory task under several levels of difficulty and rated the workload associated with each using the Subjective Workload Assessment Technique (SWAT). SWAT ratings proved sensitive to two of

the three difficulty manipulations in the memory task, and demonstrated greater sensitivity in this respect than either of two primary task measures that were employed. The results extend the applicability of SWAT to the type of spatial memory task used and, therefore, provide further support for the general applicability of SWAT as a workload measurement technique.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 2, p. 680. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

061 Eggleston, R. G., & Kulwicki, P. V. (1984). A technology forecasting and assessment method for evaluating system utility and operator workload. Proceedings of the Human Factors Society 28th Annual Meeting, 1, 31-35.

In the context of system design, technology assessment often proceeds in a bottom-up fashion, beginning with an evaluation of individual candidate system technologies. Issues of technology integration and human factors considerations are not addressed until later in system development. This paper proposes the use of a top down approach to technology assessment. The top-down approach has the advantage of evaluating technologies along dimensions of system utility, as well as technical performance, and treating human factors early in conceptual design. A technology forecasting and assessment methodology is described wherein technologies can be evaluated quantitatively in terms of technical performance, system utility, and operator workload.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 1, p. 31. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

062 Eggleston, R. G., & Quinn, T. J. (1984). A preliminary evaluation of a projective workload assessment procedure. Proceedings of the Human Factors Society 28th Annual Meeting, 2, 695-699.

Operator workload often is given little consideration early in the design of a new system. Here we describe a projective application of the Subjective Workload Assessment Technique that can be used to evaluate workload implications of technology options before they exist in hardware form. Features of the technique that probably impact its reliability and validity are discussed. Further, a comparison of projected and measured workload was made for a set of system technologies. The results indicate that the projective workload assessment technique may be able to provide a meaningful workload analysis early in system design.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 2, p. 695. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

063 Ellison, M. G., & Roberts, B. B. (1985). Timebased analysis of significant coordinated operations (TASCO): A cockpit workload analysis technique. Proceedings of the Human Factors Society 29th Annual Meeting, 2, 774-778.

TASCO has been developed to provide a diagnostic tool to aid in avionics operation task structuring. The objective of the TASCO logic is to determine the optimum organization of cockpit activities considering task complexity and task execution time compared to estimated time available to perform the task set. The TASCO cockpit workload analysis technique measures and evaluates the relationships between pilot proficiency, experience, and weapon system complexity in order to reduce risk in task performance.

(From Proceedings of the Human Factors Society 29th Annual Meeting, 1985, 2, p. 774. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

064 Ephrath, A. R., Tole, J. R., Stephens, A. T., & Young, L. R. (1980). Instrument scan - Is it an indicator of the pilot's workload? Proceedings of the Human Factors Society 24th Annual Meeting, pp. 257-258.

This presentation describes an investigation of the relationship between an aircraft pilot's visual scanning of instruments and his level of mental activity during a simulated approach and landing. The study was motivated by the increasing concern, in several areas of man-machine interaction, with the effects of changes in manual control and monitoring procedures on mental workload. This concern is particularly keen with regard to airline pilots, air traffic controllers, power plant operators, and personnel in control of large ocean-going vessels, since the cost of error can be quite high in any of these man-machine systems.

(From Proceedings of the Human Factors Society 24th Annual Meeting, 1980, p. 257. Copyright 1980 by the Human Factors Society, Inc. Reprinted by permission.)

065 Fisk, A. D., Derrick, W. L., & Schneider, W. (1983). The assessment of workload: Dual task methodology. Proceedings of the Human Factors Society 27th Annual Meeting, 1, 229-233.

The present paper outlines three major assumptions often implicitly assumed in dual task experiments conducted to assess operator workload. These assumptions are shown to be incorrect. Three criteria which should be met in dual task experiments that draw inferences from secondary task decrements are discussed.

An experiment, meeting the proposed criteria, was conducted which demonstrated that when the criteria are met secondary task performance can be predictive of primary task difficulty. However, the data also indicate that a simple assessment of effort alone will not predict total task performance.

(From Proceedings of the Human Factors Society 27th Annual Meeting, 1983, 1, p. 229. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

066 Frazier, M. L., & Crombie, R. B. (Eds.). (1982). Proceedings of the workshop on flight testing to identify pilot workload and pilot dynamics (AFFTC-TR-82-5). Edwards Air Force Base, CA: Air Force Flight Test Center. (DTIC No. AD-A129 333)

These proceedings contain more than 40 papers or abstracts of presentations made at the workshop held at Edwards AFB, California, on 19-21 January 1982. Section I contains overview papers on the subjects of pilot workload, pilot dynamics, and flight test requirements. Included in this section are papers describing the various mission environments encountered by advanced aircraft and design procedures used to improve pilot-vehicle performance capability. Section II addresses the quantification of pilot workload using measures of spare mental capacity, subjective ratings and opinion, and pilot physiology. Section III covers the identification of pilot dynamics and task performance including the measurement of pilot performance, the modeling of pilot dynamics, and the collection of flight test data. General conclusions of the workshop were that workload is a multidimensional concept involving complex relationships among the pilot, the vehicle, the tasks, and the environment.

067 Gill, R. T., & Wickens, C. D. (1983). Operator workload as a function of the system state: An analysis based upon the event-related brain potential. Proceedings of the 18th Annual Conference on Manual Control, AFWAL-TR-83-3021, pp. 100-107. (DTIC No. AD-A131 256)

While the processing demands of second order manual control are known to be greater than those of 1st or 0 order, the precise nature of locus of these increased demands is not well established. The purpose of this research is to determine if the demands are perceptual--related to the perception of higher derivatives of the error signal or characteristics of the system state, and thereby fluctuating with changes in these variables or central. In the latter case, we assume the demand to be constant over time, a consequence of the increased demands of activating a more complex internal model. Event-related brain potentials (ERP) -- more specifically, P300 amplitude -- were employed to assess operator workload while controlling a second order system. The ERP waveforms were categorized according to the system state at the time of the eliciting probes. Statistical analyses revealed no differences in P300 amplitude among the categories. Thus, it was concluded that the increased level of operator workload remained constant rather than fluctuating with changes in the system state. These results identify central processing rather than perception as the locus of higher order load.

068 Glenn, F. A., Zaklad, A. L., & Wherry, R. J., Jr. (1982). Human operator simulation in the cognitive domain. Proceedings of the Human Factors Society 26th Annual Meeting, pp. 964-968.

This paper outlines and briefly discusses some major results of the Human Operator Simulator (HOS) development program. HOS is described as it is currently configured, and some thoughts are presented on the further development of HOS in the cognitive domain.

(From Proceedings of the Human Factors Society 26th Annual Meeting, 1982, p. 964. Copyright 1982 by the Human Factors Society, Inc. Reprinted by permission.)

069 Goettl, B. P. (1985). The interfering effects of processing code on visual memory. Proceedings of the Human Factors Society 29th Annual Meeting, 1, 66-70.

The possibility of reducing task interference in complex aviation environments by taking advantage of the verbal-spatial short term memory dichotomy is explored in a dual-task paradigm. Eighteen subjects performed verbal and spatial retention memory tasks concurrently with intervening verbal and spatial cognitive tasks. Both number and processing code of the intervening tasks were manipulated. Support was found for code-specific interference such that concurrently performed tasks of the same code disrupted performance more than concurrent tasks of different codes. In addition spatial memory was found to be more fragile than verbal memory. Implications of the findings to mental workload reduction include task scheduling, presentation format, and assignment.

(From Proceedings of the Human Factors Society 29th Annual Meeting, 1985, 1, p. 66. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

070 Gopher, D. (1980). On the training of time sharing skills: An attention viewpoint. Proceedings of the Human Factors Society 24th Annual Meeting, pp. 259-263.

Training of time-sharing skills is discussed within an attention framework in which poor time-sharing performance is interpreted to stem from scarcity or inefficient utilization of processing resources. Practice is argued to increase resource availability either by reducing the resource demands of each task, improving coordination, or enhancing the voluntary control on resource allocation. Based on this analysis notions of skill generalizations and implications for the development of training procedures are examined.

(From Proceedings of the Human Factors Society 24th Annual Meeting, 1980, p. 259. Copyright 1980 by the Human Factors Society, Inc. Reprinted by permission.)

071 Gopher, D. (1984). Measurement of workload: Physics, psychophysics, and metaphysics. Proceedings of the 20th Annual Conference on Manual Control, NASA CP-2341, 2, 55 (abstract only).

The measurement of operator workload is an issue of great concern in the design and evaluation of modern engineering systems. This concern has led to the development of a wide arsenal of measurement techniques, all intended to quantify the phenomena accompanying the behavior of the human processing system when its capacity to meet task demands has been exceeded. Three general categories of measurement approaches are, performance based measures, physiological indices, and subjective scales. In theory, the three approaches should constitute alternative strategies to expose the hidden limitations of internal processors. In practice, there is only a sparse knowledge on the relationship between workload measures obtained under different approaches. Moreover, there appears to be a debate among proponents of these approaches on the validity, comprehensiveness and exclusiveness of different measures. The present paper reviews the results of two experiments in which workload analysis was conducted based upon performance measures, brain evoked potentials and magnitude estimations of subjective load. The three types of measures were jointly applied to the description of the behavior of subjects in a wide battery of experimental tasks. Data analysis shows both instances of association and dissociation between types of measures. A general conceptual framework and methodological guidelines are proposed to account for these findings.

072 Gopher, D. (1984). Workload book: Assessment of operator workload in engineering systems (NASA CR-166596). Washington, DC: National Aeronautics and Space Administration.

The report describes the structure and initial work performed toward the creation of a handbook for workload analysis directed at the operational community of engineers and human-factors psychologists. The goal of the report, when complete, will be to make accessible to such individuals the results of theoretically-based research that are of practical interest and utility in the analysis and prediction of operator workload in advanced and existing systems. In addition, the results of a laboratory study focused on the development of a subjective rating technique for workload that is based on psychophysical scaling techniques are described.

073 Gopher, D, & Braune, R. (1984). On the psychophysics of workload: Why bother with subjective measures? Human Factors, 26, 519-532.

Psychophysical functions describe the relationship between variations in the amplitude of a defined physical quantity and the psychological perception of these changes. Examples are brightness, loudness, and pain. The regularities of these relationships have been formulated into psychophysical laws. The measurement methodology of psychophysical scaling has been refined by the Harvard group led by S. S. Stevens, who proposed a power function as a general form for such laws. The main argument of the present article is that a similar scaling approach can be adapted to the measurement of workload and task demands based upon subjective estimates. The rationale is that these estimates, like other psychophysical judgments, reflect the individual's perception of the amount of

processing resources that the subject invests to meet the demand imposed by a task. This approach was successfully applied to the assessment of 21 experimental conditions given to a group of 60 subjects. The paper discusses the main results of this effort and their implications to theory and application in human performance.

(From Human Factors, 1984, 26, p. 519. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.

074 Gopher, D., Chillag, N., & Arzi, N. (1985). The influence of voluntary effort, context, and anchor task on the subjective estimate of load (Final Report for NASA Grant NAGW-494). Haifa, Israel: Technion.

Subjects (Ss) performed two different experimental tasks during three separate trials. The two tasks involved either size or letter matching and were performed individually for two trials (single task) and combined for the last (dual task). During the experiment, Ss received continuous feedback. Load was increased by varying the precision required for making an appropriate response. Immediately after each trial, Ss were asked to report the extent of load they experienced. Ss' perceptions of greater workload were related to higher precision and dual-task conditions. However, except for performance fluctuations for the secondary task in the dual-task condition, primary task performance remained constant.

075 Gopher, D., Chillag, N., & Arzi, N. (1985). The psychophysics of workload - A second look at the relationship between subjective measures and performance. Proceedings of the Human Factors Society 29th Annual Meeting, 2, 640-644.

Load estimates based upon subjective and performance indices were compared for subjects performing size matching and letter typing tasks under 6 levels of priorities, in single and dual task conditions. Each half of the group used a different task as reference in their subjective judgement. The results are interpreted to indicate that subjective measures are especially sensitive to voluntary allocation of attention and to the load on working memory. Association with performance is expected whenever these two factors are main determinants of performance efficiency, otherwise the two are likely to dissociate.

(From Proceedings of the Human Factors Society 29th Annual Meeting, 1985, 2, p. 640. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

076 Gopher, D., & Donchin, E. (1986). Workload--an examination of the concept. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), Handbook of perception and human performance: Vol. II. Cognitive processes and performance (pp. 41-1 - 41-49). New York: John Wiley & Sons.

This chapter represents a theoretical examination of the multidimensional, multifaceted concept of workload. Due to the complexity of the construct, no single measure is capable of capturing all relevant aspects, nor may multiple measures covary within a single task. The discussion was concerned with clarifying the nature of the dimensions along which workload varies to explicate

the attributes that should be considered in the selection of a measurement procedure. The primary thesis is that workload assessment focuses on measuring the processing and response limitations of the human information processing system which are revealed through the interactions between an operator and the assigned tasks. The nature of the limitations were considered on two levels: (1) the more theoretical level (in which the invariant, open loop properties of the human processing system were examined) and (2) a more practical level (in which workload was characterized, at any instant, as the joint, closed loop property of the human and the assigned task). In general the focus of the theoretical discussions emphasized the close affinity between the study of workload and attention, with an additional discussion of the energetical and structural characteristics of the central processor. The recommendation was made that measurement procedures should encompass both conscious and nonconscious processing activities; a detailed task analysis should be performed to uncover the major components of the task, followed by a battery of performance-based measures designed to evaluate the load on each component.

077 Gopher, D., & North, R. A. (1974). The measurement of attention capacity through concurrent task performance with individual difficulty levels and shifting priorities. Proceedings of the Human Factors Society 18th Annual Meeting, pp. 480-485.

Some of the unsolved problems in the application of secondary task techniques include: (a) the evaluation of relative changes in performance in dual task situations; (b) the prediction of possible interactions between different tasks and their components; and (c) the extent of voluntary control of capacity allocation. The present paper describes a three-phase experiment in which an effort was made to attack these problems by a new methodological approach. The three successive phases included separate performance of the experimental tasks (one dimensional compensatory tracking and a digit processing, reaction time task) with adaptive adjustment of difficulty, simultaneous performance of the tasks with equal task priorities, and simultaneous performance with several manipulations of the two task priorities. The results have demonstrated the usefulness of the general methodological approach for the assessment of capacity limitations as well as for the evaluation of possible interactions between tasks. With regard to the allocation of capacity, the experimental results proved that, in general, subjects were able to adjust their allocation of capacity to the various changes in task priorities.

(From Proceedings of the Human Factors Society 18th Annual Meeting, 1974, p. 480. Copyright 1974 by the Human Factors Society, Inc. Reprinted by permission.)

078 Graaff, R. C. van de. (1982). NLR research on pilot dynamics and workload. In M. L. Frazier & R. B. Crombie (Eds.), Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics (AFFTC-TR-82-5, pp. 79-90). Edwards Air Force Base, CA: Air Force Flight Test Center. (DTIC No. AD-A129 333).

This paper reviews some results of NLR research on pilot performance and workload both in manual control and monitoring tasks. A substantial part of the research has been devoted to the description of human control behavior and task conditions

in terms of modern control theory. The principal results of this modelling approach are considered in the first part of this paper. It is shown that the modelling approach can provide an adequate framework for analysis of pilot dynamics and workload. The second approach, discussed in the subsequent part of this paper, deals with the assessment of pilot workload by means of physiological parameters. The results of an in-flight experiment reflect a good correlation between measures of heart rate and respiration frequency, model results, subjective ratings, overall performance and control activities.

079 Gulick, R. (1982). Validation of pilot workload estimates utilizing in-flight data. In M. L. Frazier & R. B. Crombie (Eds.), Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics (AFFTC-TR-82-5, pp. 254-274). Edwards Air Force Base, CA: Air Force Flight Test Center. (DTIC No. AD-A129 333)

The purpose of this briefing is to describe the current capabilities of the Douglas Aircraft Company's workload evaluation techniques. The objectives of our workload evaluation studies are to isolate design deficiencies and operational problems associated with specific hardware designs and functional arrangements of equipment during initial stages of development, to verify that crew duty allocations and operational procedures do not exceed the capabilities of individual crewmen, and to evaluate crew composition. The study approach includes the development of design mission scenarios and an indepth evaluation of selected mission segments using task/timeline analytical techniques. Outputs of the crew workload evaluation program include: task loadings of individual crew members, equipment interface, and body channels. Outputs also include external vision availability and information processing time.

080 Gunning, D., & Manning, M. (1980). The measurement of aircrew task loading during operational flights. Proceedings of the Human Factors Society 24th Annual Meeting, pp. 249-252.

In-flight observers collected task loading data during four operational flights of the Air Force's KC-135 tanker aircraft. The data were collected via a Datamyte Data Recorder which the observers used to record the occurrences of 45 different tasks performed by the pilot, copilot, and navigator. Examples of the task loading data are presented including workload profiles for each crewmember, peak workload situations, and usage data for individual aircraft systems.

(From Proceedings of the Human Factors Society 24th Annual Meeting, 1980, p. 249. Copyright 1980 by the Human Factors Society, Inc. Reprinted by permission.)

081 Hancock, P. A. (1986). The role of temporal factors in workload prediction. Proceedings of the 1986 IEEE International Conference on Systems, Man, and Cybernetics, 2, 1049-1053.

In examining the role of time in mental workload, this paper presents a different perspective from which to view the problem of assessment. Workload is plotted in three dimensions, whose axes represent effective time for action, perceived distance from desired goal state, and level of effort required to achieve such a

goal. This representation allows the generation of isodynamic workload contours which incorporate the factor of operator competence. A simple physical analogy for this representation indicates an avenue toward quantification, and subsequently the potential for useful workload prediction.

(Copyright 1986 IEEE. Reprinted by permission from Proceedings of the 1986 IEEE International Conference on Systems, Man, and Cybernetics, Atlanta, GA, October 11-13, p. 1049.)

082 Hancock, P. A., & Chignell, M. H. (1986). Toward a theory of mental workload: Stress and adaptibility in human-machine systems. Proceedings of the 1986 IEEE International Conference on Systems, Man, and Cybernetics. 1, 378-383.

In light of the present difficulties in assessment, there is a pressing need for a general theory of mental workload (MWL). This paper explores a view of the task as a stress and highlights the commonalities between mental workload and the psychological and physiological reactions of an operator to stress in general. In the absence of a normative theory of mental work, mental workload is defined as an organismic response to the requirements of the task. The role of mental workload assessment within an adaptive human-machine system is outlined, and the use of changes in cognitive functioning to predict subsequent failure in human performance is recommended.

(Copyright 1986 IEEE. Reprinted by permission from Proceedings of the 1986 IEEE International Conference on Systems, Man, and Cybernetics, Atlanta, GA, October 11-13, p. 378.)

083 Harris, R. L., Sr., Tole, J. R., Ephrath, A. R., & Stephens, A. T. (1982). How a new instrument affects pilots' mental workload. Proceedings of the Human Factors Society 26th Annual Meeting, pp. 1010-1013.

Display evaluation has generally been performed using reaction time experiments or subjective evaluations to determine which display is better suited for a particular application. A new testing procedure combined with two analysis techniques of scanning behavior have been used to evaluate two alternative display designs in a realistic setting. Tests were conducted in the Langley Research Center's General Aviation Simulator. The results indicated that a new bargraph type of vertical speed indicator located between the attitude indicator and the altimeter is looked at more quickly, appears to lower cognitive workload slightly, and was preferred over a conventional vertical speed indicator.

(From Proceedings of the Human Factors Society 26th Annual Meeting, 1982, p. 1010. Copyright 1982 by the Human Factors Society, Inc. Reprinted by permission.)

084 Hart, S. G. (1982). Workload assessment research program. Invited address at the Air Force Office of Scientific Research in Biocybernetics and Workload Annual Review, Alexandria, VA.

The goal of this program is to develop relevant and reliable measures of pilot workload to assess and predict the impact of aircraft and ATC system changes on aircrews. Although pilots typically adjust to advances in technology, there may be unacceptable costs associated with the adjustment: pilot overload, stress or fatigue, additional training, or reduced safety. The effectiveness with which aircrews use new and existing equipment is usually defined by their performance whereas the cost to the aircrew of producing such performance is pilot workload. Measures of performance and workload may not covary, however, as pilots may or may not be willing or able to meet increased task demands. Further, existing measures of physical workload and overt performance may not reflect the cognitive and perceptual activities which are a major element in piloting current and future aircraft.

The term "workload" serves as a convenient label for a number of events, ideas, states, and dimensions. These factors may either relate to the operator or to the task, they may covary or not, and they may derive from the task at hand or simply coexist with it. There may be only one of these factors, however, that is uniquely "workload" and not something else: the operator's perception of his experience. If an operator feels loaded, then he is loaded and this will be reflected in physiological, subjective, and objective measures, although not necessarily in performance. This experience is derived from the other factors, but the importance placed on different components varies from person to person. Because workload measures typically reflect a fraction of the total situation and may not focus on dimensions that are relevant to that operator, available measures are often unreliable and uninformative.

Due to the complexities involved, many fundamental issues must be resolved before appropriate and reliable measures can be developed and applied: (1) Standardize the selection and combination of flight-related tasks so that predictable types and levels of primary task demand can be imposed; (2) determine the effects of many factors, such as task demands, fatigue, time pressure, effort, success, and the circumstances under which single or multiple tasks are performed on the perception of workload; (3) identify the effective level of task demand and effort as a function of the level of automatic processing and control; (4) determine the sensitivity and intrusiveness of commonly used workload measures; (5) analyze pilot errors and communications as primary task measures of workload; and (6) produce a practical guide for the analysis of workload.

085 Hart, S. G. (1986). The relationship between workload and training: An introduction. Proceedings of the Human Factors Society 30th Annual Meeting, 2, 1116-1120.

This paper reviews the relationships among workload, performance and training. It is intended to serve as an introduction for the remaining papers in this symposium. Its goal is to introduce the concepts of workload and training and to suggest how they may be related. It suggests some of the practical and theoretical benefits to be derived from their joint consideration: training effectiveness can be improved by monitoring trainee workload and the reliability

of workload predictions and measures can be improved by identifying and controlling the training levels of experimental subjects.

(From Proceedings of the Human Factors Society 30th Annual Meeting, 1986, 2, p. 1116. Copyright 1986 by the Human Factors Society, Inc. Reprinted by permission.)

086 Hart, S. G. (1986). Theory and measurement of human workload. In J. Zeidner (Ed.), Human productivity enhancement: Vol. I. Training and human factors in system design (pp. 396-455). New York: Praeger.

The goal of this chapter is to define human workload, what influences it, how it is measured, and why it is of theoretical and practical concern. The first section reviews typical definitions and motives for measuring and predicting workload. A structure is proposed to relate and integrate many of the factors that create or influence it (e.g., the demands imposed in a man-machine system, its response to them, and the subjective experiences of operators). A third section describes five types of assessment and predictive methodologies: (1) subjective ratings, (2) primary task performance, (3) secondary task performance, (4) physiological recordings, and (5) analytic procedures. Finally, the selection and application of appropriate tools to predict or assess imposed workload, system performance and behavior, or operator experience are considered.

(Based on material in Sandra G. Hart, "Theory and Measurement of Human Workload," in Human Productivity Enhancement, Joseph Zeidner, Ed. [Praeger Publishers, New York, a division of Greenwood Press, Inc., 1986]. Copyright 1986 by Praeger Publishers. Reprinted by permission.)

087 Hart, S. G., Battiste, V., Chesney, M. A., Ward, M. M., & McElroy, M. (in press). Type A vs. Type B: Comparison of workload, performance and cardiovascular measures. Human Factors.

This study provided experimental evidence for the hypothesized difference in cardiovascular responsiveness between Type A and B men. Further, it identified one source of variability in subjective, cardiovascular, and behavioral measures obtained in response to the same task. The responses of 14 Type A and 14 Type B males were compared in a multi-task, supervisory control simulation. Task difficulty was varied by manipulating the number of elements per task, the rate they were available for performance, and the time between successive tasks. Although the predicted relationships between experimental manipulations, performance measures, and subjective ratings were found for all subjects, the strategies adopted by Type A's and their subjective experiences were influenced by task manipulations more than they were for Type B's. Greater cardiovascular responsiveness was found for Type A subjects than for Type B; both systolic and diastolic levels increased significantly as task difficulty was increased for Type A subjects, but remained constant for Type B. Heart rate was unaffected by task difficulty manipulations and did not covary with rated workload.

088 Hart, S. G., Battiste, V., & Lester, P. T. (1984). POPCORN: A supervisory control simulation for workload and performance research. Proceedings of the 20th Annual Conference on Manual Control, NASA CP-2341, 1, 431-453.

A multi-task simulation of a semi-automatic supervisory control system was developed to provide an environment in which training, operator strategy development, failure detection and resolution, levels of automation, and operator workload can be investigated. The goal was to develop a well-defined, but realistically complex, task that would lend itself to model-based analysis. The name of the task ("POPCORN") reflects the visual display that depicts different task elements milling around waiting to be released and "pop" out to be performed. The operator's task was to complete each of 100 task elements that were represented by different symbols, by selecting a target task and entering the desired a command. The simulated automatic system then completed the selected function automatically. Task difficulty, operator behavior, and experienced workload were varied by manipulating: (1) the number of elements per task; (2) the number of discrete tasks; (3) the penalties for lagging behind the system; (4) task schedule; and (5) payoff structure for performing or failing to perform task elements. Highly significant differences in performance, strategy, and rated workload were found as a function of all experimental manipulations (except reward/penalty). In addition, a proposed technique for reducing the between-subject variability of workload ratings was described and applied successfully. The first simulation conducted with this task defined a range of scenarios that imposed distinctly different levels of workload on operators and resulted in different levels of performance and operator strategies.

089 Hart, S. G., & Bortolussi, M. R. (1984). Pilot errors as a source of workload. Human Factors, 26, 545-556.

A pilot opinion survey was conducted to develop a database for creating simulation scenarios that impose predetermined levels of pilot workload. Twelve pilots estimated the effect of 163 events and activities (which they had encountered during their previous flying experiences) on performance, effort, workload, and stress. The events, described in the context of flight scenario segments, included control, navigation and communications activities, aircraft and system failures, and pilot errors. In general, workload, stress, and effort ratings were significantly correlated with each other but not with performance ratings; however, some different response patterns were found as a function of flight segment (e.g., workload, stress, and performance, but not effort, ratings varied with flight phase) and type of event. Errors were rated as a significant source of change for workload, stress, and performance, suggesting that errors could be conceptualized as a cause of workload rather than as a symptom.

090 Hart, S. G., & Chappell, S. L. (1983). Influence of pilot workload and traffic information on pilot's situation awareness. Proceedings of the 19th Annual Conference on Manual Control, pp. 4-26.

Although it seems intuitively obvious that the addition of a cockpit display of traffic information (CDTI) should enhance a pilot's awareness of the current and projected situation of other aircraft, it has not been empirically determined that such is the case. Furthermore, there is some question about the utility of CDTI under conditions of relatively high pilot workload; when pilots become busy they may ignore the CDTI or take unilateral actions based on incompletely understood information. The current simulation was designed to determine how much information pilots could recall about eight aircraft simultaneously participating in a simulated approach task. A stop-action technique was used, so that each approach sequence was terminated at some point and the participants completed a written debriefing describing their recall of aircrafts' positions, situations, and intentions. The experimental variables included: (1) presence or absence of CDTI; (2) CTDI quality; and (3) level of concurrent workload. Four groups each consisting of three transport pilots, four instrument-rated general aviation pilots and one controller participated in the experiment. Concurrent workload but not CDTI quality or presence significantly affected the type and amount of information remembered. Rated workload and several types of communications were increased by the addition of CDTI and by the experimental manipulations intended to increase workload. The pilots reported feeling that CDTI afforded them a better understanding of the traffic situation, but this subjective impression was not supported by an improvement in the amount of information recalled.

091 Hart, S. G., Childress, M. E., & Bortolussi, M. (1981). Defining the subjective experience of workload. Proceedings of the Human Factors Society 25th Annual Meeting, pp. 527-531.

Flight scenarios that represent different types and levels of pilot workload are needed in order to conduct research about, and develop measures of, pilot workload. In order to be useful, however, the workload associated with such scenarios and the component tasks must be determined independently. An initial study designed to provide such information was conducted by asking a panel of general aviation pilots to evaluate flight-related tasks for the overall, perceptual, physical, and cognitive workload they impose. These ratings will provide the nucleus for a data base of flight-related primary tasks that have been independently rated for workload to use in workload assessment research.

092 Hart, S. G., Childress, M. E., & Hauser, J. R. (1982). Individual definitions of the term "workload." Proceedings of the Eighth Psychology in the DOD Symposium, pp. 478-485.

A study was conducted in which four groups of raters (51 researchers, 28 college students, 12 general aviation pilots, and 26 high school students) assigned 19 possible components of workload to one of three categories: (1) not related to workload; (2) related to, but not a primary component of workload; and (3) a primary element of workload. These ratings were factored to determine the relationships among the items. The analysis yielded seven factors: fatigue/stress, task difficulty, effort, performance/motivation, task type, interest in task, and purpose of task. The 117 participants were clustered on the within-subject standardized factor scores. This analysis yielded seven patterns of responses about the relative primacy of the different factors to different individuals' definitions of workload. The results indicate that patterns of estimating the primacy of components in subjective workload evaluation exist which cross working group lines.

093 Hart, S. G., Hauser, J. R., & Lester, P. T. (1984). Inflight evaluation of four measures of pilot workload. Proceedings of the Human Factors Society 28th Annual Meeting, 2, 945-949.

Four measures of pilot workload were tested in the NASA C-141 Kuiper Airborne Observatory. The measures included a communications analysis, subjective ratings of workload, subjective ratings of additional factors related to workload, and heart rate. Data were collected for 11 flights, each of which lasted approximately seven hours. Heart rate was found to be significantly higher for the pilot who was flying than for the pilot who was not flying and it varied significantly across flight segments, peaking during landing and take off, particularly for the pilot in the left seat who was responsible for aircraft control. For both left and right seats, the subjective assessment of stress rather than the subjective assessment of workload was significantly correlated with variation in heart rate. Frequencies of different types of communications varied significantly across segments of flight, however, they were not correlated with subjective ratings of workload. There was a significant difference between the left and right seats in the types of activities that contributed to their workload, however, workload was considered to be equivalent for the two.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 2, p. 945. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

094 Hart, S. G., Sellers, J. J., & Guthart, G. (1984). The impact of response selection and response execution difficulty on the subjective experience of workload. Proceedings of the Human Factors Society 28th Annual Meeting, 2, 732-736.

The influence of variations in response selection and response execution difficulty on the workload and performance of 11 experimental subjects was investigated. The 20 laboratory tasks they performed involved a binary response selection that required different levels of mental processing (e.g., choice

reaction time, prediction, memory search, etc.). A target acquisition task was added following response selection on half of the trials. A weighted combination of bipolar ratings on nine workload-related dimensions was used to evaluate the workload experienced by the subjects. In addition, subjects rank-ordered the tasks with respect to workload before (a prediction) and again after (a retrospective comparison) performing them. Apparently minor variations in stimulus presentation resulted in significantly increased reaction times and workload ratings, as did the more obvious manipulations of response selection load. The addition of the target acquisition task increased workload ratings and reaction times, however the "cost" of performing the two-stage task (as indicated by measures of speed, accuracy, and subjective opinion) was considerably less than would be expected by combining measures for the component tasks. Movement times for the target acquisition tasks increased significantly as a function of the index of difficulty of the target, but were not affected by the difficulty of the response selection task.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 2, p. 732. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

095 Hart, S. G., & Sheridan, T. B. (1984). Pilot workload, performance, and aircraft control automation. Human Factors Considerations in High Performance Aircraft (AGARD CP-371, pp. 18.1-18.12). Neuilly-Sur-Seine, France: North Atlantic Treaty Organization - Advisory Group for Aerospace Research and Development.

This report reviews conceptual and practical issues associated with the design, operation, and performance of advanced systems and the impact of such systems on the human operators. The development of highly automated systems has been driven by the availability of new technology and the requirement that operators safely and economically perform more and more activities in increasingly difficult and hostile environments. It has become obvious that the workload of the operators, particularly their mental workload, may become a major area of concern in future design considerations. There has been, however, little research to determine how automation and workload relate to each other, although it is assumed that the abstract, supervisory, or management roles that are assumed by operators of highly automated systems will impose increased mental workload. The relationship between performance and workload, which is poorly understood at best for relatively simple tasks, will be discussed in relation to highly complex and automated environments.

(The original version of this material was first published by the Advisory Group for Aerospace Research and Development, North Atlantic Treaty Organization [AGARD/NATO] in CP 371, Human Factors Considerations in High Performance Aircraft, 1984. Reprinted by permission.)

096 Hart, S. G., Shively, R. J., Vidulich, M. A., & Miller, R. C. (1986). The effects of stimulus modality and task integrality: Predicting dual-task performance and workload from single-task levels. Proceedings of the 21st Annual Conference on Manual Control, NASA CP-2428, pp. 5.1-5.18.

The influence of stimulus modality and task difficulty on workload and performance was investigated in the current study. The goal was to quantify the "cost" (in terms of response time and experienced workload) incurred when essentially serial task components shared common elements (e.g., the response to one initiated the other) which could be accomplished in parallel. The experimental tasks were based on the "Fittsberg" paradigm; the solution to a SternBERG-type memory task determines which of two identical FITTS targets are acquired. Previous research suggested that such functionally integrated "dual" tasks are performed with substantially less workload and faster response times than would be predicted by summing single-task components when both are presented in the same stimulus modality (visual). In the current study, the physical integration of task elements was varied (although their functional relationship remained the same) to determine whether dual-task facilitation would persist if task components were presented in different sensory modalities. Again, it was found that the cost of performing the two-stage task was considerably less than the sum of component single-task levels when both were presented visually. Less facilitation was found when task elements were presented in different sensory modalities. These results suggest the importance of distinguishing between concurrent tasks that compete for limited resources from those that beneficially share common resources when selecting the stimulus modalities for information displays.

097 Hartman, B. O. (Ed.). (1978). Methods to assess workload (AGARD-CP-216). Neuilly-Sur-Seine, France: North Atlantic Treaty Organization - Advisory Group for Aerospace Research and Development. (DTIC No. AD-A057 835)

For the measurement of workload a wide variety of methods is employed, frequently in an interdisciplinary setting. This multifaceted approach to the problems of the methods best suited to assess workload is highlighted by the range of topics covered in the meeting. The conference was organized around methodology questions and involved basic issues like measurement sensitivity, reliability and validity; instrumentation and associated techniques; and study environment ranging from the laboratory to simulator and in-flight studies.

(The original version of this material was first published by the Advisory Group for Aerospace Research and Development, North Atlantic Treaty Organizaiton [AGARD/NATO] in CP 216, Methods to Assess Workload, 1978. Reprinted by permission.)

098 Hartman, B. O., & McKenzie, R. E. (Eds.). (1979). Survey of methods to assess workload (AGARD-AG-246). Neuilly-Sur-Seine, France: North Atlantic Treaty Organization - Advisory Group for Aerospace Research and Development (DTIC No. AD-A078 319)

Contents: Concepts of Workload; Concepts of Fatigue; Concepts of Stress; Some Considerations Concerning Methods to Evaluate and Assess Workload in Aircraft Pilots; Physiologic Aspects of Workload/Fatigue/Stress; Some Insights Relative to the Man-Machine System--an Overview of Ten Years of Research; Aircrew Workload Assessment Techniques; Workload Assessment Methodology Development; Quantitative Military Workload Analysis; Visual Performance--A Method to Assess Workload in the Flight Environment; Handling Qualities, Workload, and Heart Rate; Brain Waves and the Enhancement of Pilot Performance; Pupillometric Methods of Workload Evaluation: Present Status and Future Possibilities; Aircrew Performance Research Opportunities Using the Air Combat Maneuvering Range (ACMR); Speech Patterns and Aircrew Workload; An Exploratory Study of Psychophysiological Measurement as Indicators of Air Traffic Control Sector Workload; Individual and System Performance Indices for the Air Traffic Control System; Workload and Stress in Air Traffic Controllers; and Assessment Correlates of Workload and Performance.

(The original version of this material was first published by the Advisory Group for Aerospace Research and Development, North Atlantic Treaty Organization [AGARD/NATO] in AG 246, Survey of Methods to Assess Workload, 1979. Reprinted by permission.)

099 Hauser, J. R., Childress, M. E., & Hart, S. G. (1983). Rating consistency and component salience in subjective workload estimation. Proceedings of the 18th Annual Conference on Manual Control, AFWAL-TR-83-3021, pp. 127-149. (DTIC No. AD-A131 256)

Twelve general aviation pilots participated in a two-day experiment performing four tasks intended to load on different cognitive, perceptual, and motor dimensions. The tasks were varied in apparent difficulty level so that each pilot performed a total of sixteen tasks counter-balanced for task and level. Subjective ratings of factors contributing to workload were made immediately following each level of each task using a 15 bipolar adjective scale. Results indicated that the subjective perception of workload was not related to actual performance measures; however, the subjective ratings were generally consistent with the demands made by the levels of each task. Although only two of the rating scale items, own Performance and Task Difficulty, demonstrated significant within-task differences for all four tasks, the majority of rating scales showed within-task differences for those tasks that imposed higher cognitive demands. Strong relationships were found between Overall Workload, Stress Level, and Task Difficulty ratings on all tasks.

100 Hauser, J. R., & Hart, S. G. (1983). The effect of feedback on subjective and objective measures of workload and performance. Proceedings of the Human Factors Society 27th Annual Meeting, 1, 144 (abstract only).

Thirty subjects were employed in a mixed experimental design that examined five levels of feedback and two levels of difficulty for two tasks, with repeated measures on the difficulty and task variables. The amount and type of feedback was varied so that it provided information about performance of the task on the objective measures within a block of trials, or provided the same information at the end of the task (simply providing knowledge of results), and was also varied in quality, either as a comparison to the subject's own average performance of the task, or in comparison to an experimentally determined 'figure of merit.' Two tasks, each with two levels of difficulty, were used: (1) a task that primarily imposed cognitive demands, a version of the Sternberg memory task, and (2) a task that primarily imposed psychomotor demands, a target acquisition task modelled on the Fitts' Law paradigm. Both objective and subjective measures demonstrated reliable and predictable effects for the difficulty levels of the two tasks, however the tasks were differentially affected by the feedback conditions, but differences between and within tasks were generally small. The relationships between objective measures, and subjective ratings of workload and performance rarely reached significant levels.

(From Proceedings of the Human Factors Society 27th Annual Meeting, 1983, 1, p. 144. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

101 Hauser, J. R., & Hart, S. G. (1983). Subjective workload experienced during pursuit tracking as a function of available information. Proceedings of the 19th Annual Conference on Manual Control (p. 3). Cambridge, MA: Massachusetts Institute of Technology.

Twelve general aviation pilots performed a pursuit tracking task where the objective was to "pilot" the pursuit vehicle (a simplified delta wing aircraft) after a target (represented by a cross). Successful acquisition of the target was always displayed on the screen. Each subject experienced twenty 10-min experimental runs in a partially counterbalanced order. The experimental variables included: (1) number of dimensions (2 or 3), (2) target path complexity (low or high), and (3) availability of information (both target and pursuit vehicle were displayed for either 100%, 50%, 25% of the time, or at subject command). Subjects controlled the vehicle by pressing rocker arm switches for the functions of yaw (left and right), roll (left and right), speed (acceleration and deceleration), pitch (up and down), and screen illumination (for the one condition where subject control was given to display time). After every experimental trial, subjects rated the preceding experience using a set of 15 bipolar adjective scales. Significant differences were found for their difficulty and display time variables on the majority of the scales, and for the dimension variable on eight of the scales. Strong relationships were found between actual time-on-target with ratings of overall workload, performance, and difficulty for almost all subjects, who were also able to estimate their time-on-target with a high degree of accuracy.

102 Hawkins, H. L., & Ketchum, R. D. (1980). The case against secondary task analyses of mental workload (TR-6). Eugene: Oregon University. (DTIC No. AD-A080 792)

In a commonly used sense, mental workload refers to the proportion of an individual's total processing capacity taken up by a particular cognitive task or task combination. One approach to the assessment of mental workload is called the secondary task analysis. In this approach, the operator is required to carry out two simultaneous tasks, assigning one (the primary task) a high priority and the other (the secondary task) a lower priority. The primary task's mental workload is defined in terms of the degradation in secondary task performance occurring under dual-relative to single-task conditions. The validity of this approach critically hinges to the validity of the assumptions (a) that human processing capacity is unitary or undifferentiated; (b) that the human information processing system contains no significant task-specific capacities; and (c) that overall capacity remains invariant across changes in processing demand. The literature pertaining to these assumptions is reviewed. It is found that while many of the theoretical issues surrounding the assumptions remain unresolved, the available data argue strongly against the general advisability of the secondary task approach. The problem is that the workload ordering obtained by this approach for any set of (primary) tasks can be expected to vary with the secondary task used. Consequently, the approach will not yield a general measure of workload demand.

(Reprinted by permission.)

103 Haworth, L. A., Bivens, C. C., & Shively, R. J. (1986). An investigation of single-piloted advanced cockpit and control configurations for nap-of-the-earth helicopter combat mission tasks. Proceedings of the 1986 Meeting of the American Helicopter Society, pp. 657-672.

A two-phase handling qualities and pilot workload investigation of single-pilot operation in the combat Nap-of-the-Earth (NOE) environment was started by the Aeroflightdynamics Directorate in October 1985. Phase one of the investigation was conducted in cooperation with the NASA Ames Research Center on the NASA Vertical Motion System (VMS) simulator, using the Advanced Digital Optical Control System (ADOCS) laws and a glass cockpit. Handling Quality Ratings (HQR) and workload ratings were recorded for NOE flight task maneuvers during single-pilot and "dual-" pilot operation. Control automation/augmentation was varied to record differences between configurations for dual- and single-pilot operation. Only one control system configuration investigated was rated satisfactory for single-pilot NOE flight due to increased attentional demands placed on the pilot.

(Reprinted by permission.)

104 Heffley, R. K. (1983). Pilot workload factors in the total pilot-vehicle-task system. Proceedings of the Human Factors Society 27th Annual Meeting, 1, 234-238.

This paper is based on a current study of pilot workload models for crucial Navy flight tasks such as the carrier landing and high-speed, low-level navigation. The objective is to construct a more rigorous and complete view of the overall pilot-vehicle-task system in order to describe how facets of pilot workload can be associated with elements of the system. The purpose of the paper is to discuss workload features in a system context as a first step to developing a more thorough workload prediction process for the design and operation of aircraft.

105 Heffley, R. K., Clement, W. F., & Jewell, W. F. (1982). Overview of work in progress on non-intrusive assessment of pilot workload and pilot dynamics. In M. L. Frazier & R. B. Crombie (Eds.), Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics (AFFTC-TR-82-5, pp. 516-550). Edwards Air Force Base, CA: Air Force Flight Test Center. (DTIC No. AD-A129 333)

The purpose of this paper is to describe one general approach to non-intrusive assessment of pilot workload and pilot dynamics, how this assessment is carried out, and recent and on-going projects which involve its application. We begin with a general closed-loop formulation of piloting technique which can include both psychomotor and cognitive aspects. Examples are given showing how this basic closed-loop structure can be applied to a variety of flight tasks with the help of several quantitative examples. We then go on to consider the major objectives of work in progress, some of the tools and concepts used in non-intrusive piloting technique assessment, and a number of the important technique assessment, and a number of the important lessons learned in applying various procedures. Finally, we summarize some of the recent and on-going projects involving non-intrusive piloting technique assessment procedures.

106 Helm, W. R. (1975). The function description inventory as a human factors tool in evaluating system effectiveness in operational environments. Proceedings of the Human Factors Society 19th Annual Meeting, pp. 206-208.

Determining the ultimate suitability of a major air weapon system through the test and evaluation (T&E) process include man-machine evaluations of complex components such as radar, electronic support measures, communications, etc. Traditionally, these subsystems have been evaluated by individual operators using as the prime basis for the evaluation, his limited experience with the subsystem test item. The evaluation results are typically narrative descriptions including indicating the degree of acceptability. This procedure lacks precise communication as well as a quantitative estimate of its reliability and validity. Since it is common that different operators will evaluate different subsystem components, the procedure prohibits integrated subsystem evaluation, that is, the

determination of how effective each subsystem is in the overall weapons system. The objective of this paper is to present a new method, the Function Description Inventory (FDI) as a tool for providing quantifiable assessment of the effectiveness of the man-machine interface, and as an aid toward integrated subsystem analysis in the total weapon system context.

(From Proceedings of the Human Factors Society 19th Annual Meeting, 1975, p. 206. Copyright 1975 by the Human Factors Society, Inc. Reprinted by permission.)

107 Hicks, T. G., & Wierwille, W. W. (1979). Comparison of five mental workload assessment procedures in a moving-base driving simulator. Human Factors, 21, 129-143.

Five methods of measuring mental workload (secondary task performance, visual occlusion, cardiac arrhythmia, subjective opinion rating scales, and primary task performance) were compared for sensitivity to changes in operator loading. Each was used to differentiate among low, medium, and high levels of workload defined in terms of the application point of crosswind gusts in a driving task. The results showed significant differences among workload levels for subjective opinion scales and primary performance measures of lateral deviation, yaw deviation, and steering reversals. A relative sensitivity estimate of these would be, from highest to lowest sensitivity, steering reversals and yaw deviation, rating scales, and lateral deviation. The techniques of occlusion, cardiac arrhythmia, and secondary task performance yielded no significant workload effect.

(From Human Factors, 1979, 21, p. 129. Copyright 1979 by the Human Factors Society, Inc. Reprinted by permission.)

108 Horst, R. L., Munson, R. C., & Ruchkin, D. S. (1984). Event-related potential indices of workload in a single task paradigm. Proceedings of the Human Factors Society 28th Annual Meeting, 2, 727-731.

Many previous studies of both behavorial and physiological correlates of cognitive workload have burdened subjects with a contrived secondary task in order to assess the workload of a primary task. The present study investigated event-related potential (ERP) indices of workload in a single task paradigm. Subjects monitored changing digital readouts for values that went "out-of-bounds." The amplitude of a long-latency positivity in the ERPs elicited by readout changes increased with the number of readouts being monitored. This effect of workload on ERPs is reported, along with plans for additional analyses to address theoretical implications.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 2, p. 727. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

109 Hutchins, C. W. (1974). A computer aided function allocation and evaluation system (CAFES). Proceedings of the Human Factors Society 18th Annual Meeting, pp. 350-353.

The Computer Aided Function Allocation and Evaluation System (CAFES) is a crew systems design support tool based on human engineering methodology, computer aids, human performance data and a data management system. It is intended to support crew systems engineers in man-machine research, requirements analysis, design, test, training and maintenance systems development. The system of computer models comprising the CAFES system parallels the system development cycle and provides the required level of detail consistent with a given stage of system development. The principle objectives of the present development program are to facilitate timely application of all essential elements of human factors technology in systems development, using automatic data processing techniques for rapid analysis and evaluation of crew subsystem performance as it effects total systems effectiveness.

(From Proceedings of the Human Factors Society 18th Annual Meeting, 1974, p. 350. Copyright 1974 by the Human Factors Society, Inc. Reprinted by permission.)

110 Jensen, R. S., & Chappell, S. L. (1984). Pilot performance and workload assessment: An analysis of pilot errors (Final Report for NASA Grant NAS 2-184). Moffett Field, CA: National Aeronautics and Space Administration, Ames Research Center.

The preceding taxonomy of pilot errors provides a useful tool for the human factors investigator seeking answers to basic and applied problems in a real world aviation environment. The design of simulation scenarios that impose predictable and objectively determined levels of workload on pilots is essential in analyzing aircraft systems and procedures in applied environments as well as in developing metrics of pilot workload and performance in the laboratory. The occurrence of unplanned events (such as pilot errors) during the execution of the most carefully designed simulation scenario can result in the loss of costly and important data in such experiments.

By considering errors as a source of workload rather than as a symptom or product of workload, errors may be analytically and theoretically related to experimentally controlled variations in input load. Thus, the contribution of errors to flight task scenario workload can be computed and added to the original prediction of imposed load (Hart, 1983). Whenever pilots slip, blunder, err, or even hesitate, additional workload may be created because this forces them out of well-learned, automatic sequences of actions, and requires additional effort to discover, diagnose, and resolve the consequences of the error

The belief that increased errors reflect increased workload is often expressed but less often supported by experimental results and needs clarification. The categorization of pilot-related behaviors with respect to impact on pilot workload provides a useful organizational scheme for a taxonomy of pilot behavior, with a particular emphasis on pilot errors. Such a taxonomy could be used to structure summarization and analysis of errors that are observed in flight-related research, and in reporting them in a standardized way. Errors observed

under a variety of well-defined experimental situations and summarized in a common format as shown above, provide an understanding of the degree to which variation in imposed task demands and pilot effort cause errors.

111 Jex, H. R. (1982). Measuring aircrew workload: Problems, progress, and promises. In M. L. Frazier & R. B. Crombie (Eds.), Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics (AFFTC-TR-82-5, pp. 216-221). Edwards Air Force Base, CA: Air Force Flight Test Center. (DTIC No. AD-A129 333)

An overviewing of the problems of defining, quantifying, and measuring mental workload during aircrew tasks is given based on our work in the areas of aircraft handling qualities, pilot model measurement and prediction, multi-display scanning and psychophysiological correlates of workload. The continued promise and problems with psychophysiological measures is assessed and the importance of some new multidimensional workload rating techniques is emphasized. The lack of unifying theoretical approach is identified as the main impediment to progress, and an approach is suggested, that can handle both continuous and discrete task loads. A review is given of some new workload measurement concepts such as Non-invasive Pilot Identification Program, the imbedded surrogate auxiliary task method, and the measurement of workload margin via the Cross-Coupled-Instability Task (CCIT).

112 Johnson, D. F., & Haygood, R. C. (1984). The use of secondary tasks in adaptive training. Human Factors, 26, 105-108.

The purpose of this study was to demonstrate the feasibility of conducting adaptive training by adapting primary-task difficulty on the basis of secondary-task performance. Sixteen subjects were adapted on the basis of primary-task performance, and 16 on secondary-task performance. These were matched by 32 fixed-training subjects, who were individually yoked to the adaptive-training subjects. Performance of subjects adapted on primary-task performance was superior to that of subjects adapted on secondary-task performance. However, the results show clearly that adaptation on secondary-task performance is practical and potentially useful. As in most previous studies, adaptive training was superior to fixed training.

(From Human Factors, 1984, 26, p. 105. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

113 Kantowitz, B. H. (1986). A theoretical approach to measuring pilot workload (Final Report for NASA Grant NCC 2-228). Washington, DC: National Aeronautics and Space Administration.

This final report covers efforts which can be grouped into three categories. First, and most important, theoretical advances aimed at integrating the concepts of attention and workload. Second, are empirical studies, primarily performed at Ames Research Center, that studied objective measures of pilot workload. Third, systems software written for data collection and analysis of workload experiments.

114 Kantowitz, B. H., Hart, S. G., & Bortolussi, M. R. (1983). Measuring pilot workload in a moving-base simulator: I. Asynchronous secondary choice-reaction task. Proceedings of the Human Factors Society 27th Annual Meeting, 1, 319-322.

The de facto method for measuring airplane pilot workload is based upon subjective ratings. While researchers agree that such subjective data should be bolstered by using objective behavioral measures, results to date have been mixed. No clear objective technique has surfaced as the metric of choice. We believe this difficulty is in part due to neglect of theoretical work in psychology that predicts some of the difficulties that are inherent in a futile search for the one and only best secondary task to measure workload. An initial study that used both subjective ratings and an asynchronous choice-reaction secondary task was conducted to determine if such a secondary task could indeed meet the methodological constraints imposed by current theories of attention. Two variants of a flight scenario were combined with two levels of the secondary task. Appropriate single-task control conditions were also included. Results give grounds for cautious optimism but indicate that future research should use synchronous secondary tasks where possible.

115 Kantowitz, B. H., Hart, S. G., Bortolussi, M. R., Shively, R. J., & Kantowitz, S. C. (1984). Measuring pilot workload in a moving-base simulator: II. Building levels of workload. Proceedings of the 20th Annual Conference on Manual Control, NASA CP-2341, 2, 359-371.

Studies of mental workload conducted in flight simulators usually regard flying as a unitary task. Workload is varied by changing the mission and/or turbulence and little attempt is made to evaluate the individual workload required by a specific flight sub-task. As a first effort in this direction, we chose three levels of flight sub-task complexity and measured the mental workload associated with each by an asynchronous secondary reaction-time task and by subjective ratings. The Base level of complexity was the simplest, requiring elementary maneuvers that do not utilize all the degrees of freedom of which an aircraft, or moving-base simulator, is capable. A Base task would be maintaining constant airspeed, heading or altitude. A Paired level task required two Base tasks performed simultaneously. A Complex level task required three Base tasks performed simultaneously. Primary task (flying) performance was not adversely affected by the addition of the auditory secondary reaction-time task. Rate of transmitted information (bits/sec) on the secondary task was able to discriminate among all three levels of the flight task for dual-task conditions. Furthermore, single-stimulation transmitted information rate was reliably greater than any dual-task performance, indicating that even the elementary maneuvers of the Base level imposed some mental workload. Subjective ratings also discriminated among the three levels of the flight task and in addition were sometimes able to discriminate between tasks within the same level of complexity.

116 Kantowitz, B. H., & Weldon, M. (1985). On scaling performance operating characteristics: Caveat emptor. Human Factors, 27, 531-547.

Problems associated with scaling and normalizing empirical performance operating characteristics (POCs) are examined. Normalization methods proposed by Wickens (1980) and by Mountford and North (1980) are critically evaluated. Computer simulations are used to generate raw-score and normalized POCs. The interpretation of transformed empirical POCs (Wickens, Mountford, and Schreiner, 1981) is shown to contain inconsistencies. The normalization techniques reviewed fail to resolve POC scaling problems. Caution must be exercised when interpreting transformed POCs.

(From Human Factors, 1985, 27, p. 531. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

117 Kirkpatrick, M., III, Malone, T. B., & Andrews, P. J. (1984). Development of an interactive microprocessor based workload evaluation model (SIMWAM). Proceedings of the Human Factors Society 28th Annual Meeting, 1, 78-80.

In an attempt to assess the extent to which human factors technology developed for air systems is applicable to ship systems problems, several task network simulation techniques were evaluated for aircraft carrier air operations. Such operations are complex and highly variable and include over 400 separate tasks performed by 18 different operators. The result of the assessment was that the available techniques were not acceptable, leading to the development of a model designated SIMWAM for simulation for workload assessment and modeling.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 1, p. 78. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

118 Klapp, S. T., Kelly, P. A., Battiste, V., & Dunbar, S. (1984). Types of tracking errors induced by concurrent secondary manual task. Proceedings of the 20th Annual Conference on Manual Control, NASA CP-2341, 2, 299-304.

Future one-man helicopters may require the pilot to control flight with one hand, and simultaneously manipulate other instruments using the other hand. This report of work in progress examines the nature of errors induced in a right hand tracking task (simulating flight control) when responses are required by the left hand. The present experiment focused on detection of hesitations in which the tracking joy stick remained motionless for 1/3 sec or longer.

119 Kramer, A. F., & Wickens, C. D. (1985). Event-related brain potentials and resource allocation: From dual-task decrements to dual-task integrality. Proceedings of the Human Factors Society 29th Annual Meeting, 2, 966-970.

The use of event-related brain potentials (ERP) in the study of mental workload and resource allocation is discussed. A series of studies are reviewed which, taken together, suggest that the P300 component of the ERP provides a reliable measure of the perceptual/central processing demands of a task. Furthermore, the

use of P300 in the assessment of mental workload offers the advantage of not requiring an overt response, thereby eliminating the possibility of secondary task intrusion into primary task performance. The concept of dual-task integrality is introduced and the resource demands of integral task pairs are inferred from measures of P300 amplitude. The findings are discussed in terms of resource models of mental workload.

(From Proceedings of the Human Factors Society 29th Annual Meeting, 1985, 2, p. 966. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

120 Kramer, A. F., Wickens, C. D., & Donchin, E. (1983). An analysis of the processing requirements of a complex perceptual-motor task. Human Factors, 25, 597-621.

Current concerns in the assessment of mental workload are discussed, and the event-related brain potential (ERP) is introduced as a promising mental-workload index. Subjects participated in a series of studies in which they were required to perform a target acquisition task while also covertly counting either auditory or visual probes. The effects of several task-difficulty manipulations on the P300 component of the ERP elicited by the counted stimulus probes were investigated. With sufficiently practiced subjects the amplitude of the P300 was found to decrease with increases in task difficulty. The second experiment also provided evidence that the P300 is selectively sensitive to task-relevant attributes. A third experiment demonstrated a convergence in the amplitude of the P300s elicited in the simple and difficult versions of the tracking task. The amplitude of the P300 was also found to covary with the measures of tracking performance. The results of the series of three experiments illustrate the sensitivity of the P300 to the processing requirements of a complex target acquisition task. The findings are discussed in terms of the multidimensional nature of processing resources.

(From Human Factors, 1983, 25, p. 597. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

121 Levison, W. H. (1982). The optimal control model for the human operator. Theory, validation, and application. In M. L. Frazier & R. B. Crombie (Eds.), Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics (AFFTC-TR-82-5, pp. 551-579). Edwards Air Force Base, CA: Air Force Flight Test Center. (DTIC No. AD-A129 333)

The Optimal Control Model for the Human Operator is reviewed. Underlying concepts are presented and the model structure is described. Special emphasis is given to the treatment of attentional workload, including the relationship between workload and closed-loop performance. Validating experimental results are presented, and both predictive and diagnostic applications of the model are reviewed. Areas of further model development and application are summarized.

122 Lindholm, E., Cheatham, C., & Buckland, G. (1981). Physiological and dual task assessment of workload during tracking and simulated flight (TR-82-0714). Tempe: Arizona State University. (DTIC No. AD-A119 218)

A visuomotor task of moderate complexity (tracking) and one of high complexity (simulated aircraft carrier landing) were performed alone, then in combination with a tone discrimination task at two levels of difficulty in usual dual task fashion. Measures of autonomic nervous system activation (heart rate, skin conductance) and central nervous system information processing (event related potentials) were quantified continuously during performance of all tasks. The dual task results were typical, given that most subjects treated the tone discrimination task as 'secondary' (low priority): tone discrimination performance degraded when the tone mask was combined with the tracking task and degraded even more when the tone task was conbined with the carrier landing task. While dual task methodology adequately described gross changes in workload, the physiological data permitted much more detailed interpretations and descriptions of training effects (practice), tone mask information processing, individual differences, and visuomotor task control parameters than was possible by analysis of secondary task performance. It is concluded that the physiological method has distinct advantages over the dual task method, due mostly to the nonintrusive nature and the greater detail of results afforded by the former method.

(Reprinted by permission.)

123 Lindholm, E., Cheatham, C., Koriath, J., & Longridge, T. M. (1984). Physiological assessment of aircraft pilot workload in simulated landing and simulated hostile threat environments (AFHRL-TR-83-49). Brooks Air Force Base, TX: Air Force Human Resources Laboratory. (DTIC No. AD-A140 469)

In two experiments, physiological metrics of cockpit workload were investigated in highly realistic flight simulators. In Experiment I, non-pilot males were trained on a simulated landing task and a secondary, tone discrimination task while heart rate, skin conductance, and brain event-related potentials were continuously quantified. The results showed that heart rate was a more stable measure of workload than was skin conductance. Heart rate increased during each final approach to landing, and mean heart rate decreased as the subjects gained mastery over the task as a function of practice. Four ERP components (N1, P2, N2, P3) were statistically evaluated. As workload increased, N2 became more negative and P3 became less positive; also, as workload increased, the latency difference between P3 and N1 increased. Finally, a within-subject regression analysis was employed to express the extent to which the four ERP components were intercorrelated. This measure proved to have considerable power to predict how well individual subjects would perform on the landing tasks. In Experiment 2, rated male pilots flew a simulated mission involving threat by surface-to-air missiles (SAMs). Heart rate, respiration activity, and ERPs were quantified by means of a custom-designed, miniaturized recording system. The pilots were informed of the level of SAM threat by tones sounded in the headset. The results showed that heart rate and respiration activity increased as SAM threat increased. The ERP analysis showed that N2 and P3 amplitude and P3 latency increased with threat level. The autonomic results are discussed within the

framework of activation theory, and, regarding the ERP results, it is suggested that N2 might be more important for workload and information processing studies than is P3.

124 Lyman, J. (1986). Modified Petri net model sensitivity to workload manipulations (Final Report for NASA Grant NAG 2-216). Washington, DC: National Aeronautics and Space Administration.

The purpose of this research is to investigate modified Petri nets (MPNs) as a workload modeling tool. This paper describes the results of an exploratory study of the sensitivity of MPNs to workload manipulations in a dual task. The results of the canonical correlation indicated that MPN model of the experimental task represented the task components that influenced subjective workload. Thus, the goal of this experiment was achieved by this demonstration that the MPN model was sensitive to workload changes. The next stage of this research will involve generating a classification scheme that will group events and activities that are similar in their contribution to task workload. Workload values for each class of events and activities can then be derived. This will allow testing of MPN model simulations for their prediction capability of the workload of a task.

125 Madni, A. M., & Lyman, J. (1983). Model-based estimation and prediction of task-imposed mental workload. Proceedings of the Human Factors Society 27th Annual Meeting, 1, 314-318.

Mental workload has been an area of intensive research for better than a decade. One specific area of interest in aircrew related workload research is concerned with the development of quantitative indices of workload in aircraft piloting tasks. This paper presents a model-based approach for quantifying mental workload in operational terms. The suggested modelling framework is based on an interpreted Petri net characterization of a task in which "places" are equated to specific task-related activities and "transitions" are viewed as internal forcing events. It is shown that within this framework quantitative assessments can be made of both cumulative and instantaneous workload associated with the performance of a task and its individual component subtasks. It is suggested that insights gained from analyzing task-specific workload within this modelling paradigm can suggest plausible explanations for reconciling discrepancies between subjectively elicited workload estimates and behavioral/performance measures.

(From Proceedings of the Human Factors Society 27th Annual Meeting, 1983, 1, p. 314. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

126 Mane, A., & Wickens, C. D. (1986). The effects of task difficulty and workload on training. Proceedings of the Human Factors Society 30th Annual Meeting, 2, 1124-1127.

We propose four hypotheses regarding the possible effect of workload and task difficulty on training: (1) increased levels of task difficulty will facilitate learning to the extent that these increases are (a) resource loading and (b) intrinsic to the component task to be learned. (2) Decrease of task difficulty will facilitate learning to the extent that these decreases (a) reduce the

resource load and (b) are extrinsic of the component task to be learned. (3) The learner's tendency to conserve resources may lead to the adoption of undesirable, short-term, low resource strategies early in training. (4) The effect of changes in resource demand on learning will depend upon the similarity of the resource whose demand is changed to the resource involved in learning.

(From Proceedings of the Human Factors Society 30th Annual Meeting, 1986, 2, p. 1124. Copyright 1986 by the Human Factors Society, Inc. Reprinted by permission.)

127 Mertens, H. W., Higgins, E. A., & McKenzie, J. M. (1983). Age, altitude, and workload effects on complex performance (FAA-AM-83-15). Washington, DC: Federal Aviation Administration, Office of Aviation Medicine. (DTIC No. AD-A133 594)

Fifteen healthy men in each of three age groups, 20-29 yrs, 40-49 yrs, and 60-69 yrs, were evaluated regarding complex performance in two altitude conditions (ground level vs. 3,810 m) which were administered during performance testing. Performance was measured during a 3-h test session with the Multiple Task Performance Battery (MTPB) which involved time-shared performance of several flight-related tasks presented in different combinations to vary workload. MTPB tasks consisted of monitoring of warning lights and meters, mental arithmetic, problem solving, visual target identification, and tracking. Heart rate decreased slightly at the 3,810 m altitude in the 60-69 yr group, but increased significantly at altitude in the two younger groups. Both epinephrine and norepinephrine excretion rates were highest in the 20-29 yr group and lowest in the 40-49 yr group. Age related decrements occurred in monitoring tasks, information-processing tasks, and a tracking task involving psychomotor-coordination. Performance differences occurring as a function of age were evident predominantly at moderate and high workload levels. There were no important effects of altitude on performance. Physiological and biochemical responses had little relation to performance. Implications of these findings for future research relating age to pilot performance are discussed.

128 Metzler, T. R. (1986). Register of research in progress on mental workload (AAMRL-TR-86-007). Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory. (DTIC No. AD-A168 210)

This report documents current research on operator workload. The register was compiled from responses to a questionnaire data form sent to 76 scientists who are active in basic and applied workload research. The six sections of the register include name and key-term indexes, current project descriptions, listings of workload laboratories and potential sources of research support, and a bibliography of significant publications in the area.

129 Miller, R. C., & Hart, S. G. (1984). Assessing the subjective workload of directional orientation tasks. Proceedings of the 20th Annual Conference on Manual Control, NASA CP-2341, 2, 85-95.

An experiment was conducted to investigate the impact of various flight-related tasks on the workload imposed by the requirement to compute new headings, course

changes and reciprocal headings. Eight instrument-rated pilots were presented with a series of heading-change tasks in a laboratory setting. Two levels of difficulty of each of three tasks were presented verbally (numeric values imbedded in simple commands) and spatially (headings were depicted on a graphically drawn compass). Performance was measured by evaluating the speed (response times) and accuracy (percent correct and time outs) of the responses. The workload experienced by the pilots under each experimental condition was determined by responses to a standard set of bipolar rating scales. The subjective responses and objective measures of performance reflected a strong association between subjective experience and objective behavior. The reciprocal calculations were performed quickly and accurately throughout and were considered to be minimally loading. Subjective workload, percent correct and response times for the two course-change tasks varied significantly as a function of level of difficulty and display format, with no discernable speed/accuracy trade off. The results of this study will be used to predict the workload that is imposed on pilots of actual and simulated flights by course corrections and computations in conjunction with previously obtained estimates of control and communications workload.

130 Miller, R. C., Bortolussi, M. R., & Hart, S. G. (1986). Evaluating the subjective workload of directional orientation tasks with varying displays <u>Fifth Aerospace Behavioral Engineering Technology Conference Proceedings</u> (pp. 135-138). Warrendale, PA: Society of Automotive Engineers.

An experiment was conducted to investigate the impact of various flight-related tasks on the workload imposed by the requirement to compute new headings, course changes and reciprocal headings. Nine instrument-rated pilots were presented with a series of heading-change tasks in a laboratory setting and in a single-place instrument trainer. Two levels of difficulty of each of three tasks were presented verbally (numeric values embedded in simple commands), spatially (headings were depicted on a graphically drawn compass) and combined (each of the previous displays were given simultaneously). In the instrument-trainer setting problems were presented orally by one of the experimenters and no effort was made to manipulate display types. Performance was measured by evaluating the speed (response times) and accuracy (percent correct and time outs) of the responses. The workload experienced by the pilots under each experimental condition was determined by responses to a standard set of bipolar rating scales. These subjective measures reflected the differences between levels of difficulty and types of tasks, but were generally insensitive to the manipulation of display type. The performance measures, however, displayed significant differences for all manipulations. Problems presented in the combined and alpha display formats, were done significantly faster and with significantly greater accuracy than problems in the compass format alone suggesting that the pilots were primarily using the alpha information contained in the combined display to perform the calculations. Workload ratings for the compass-only laboratory condition and the instrument trainer portion of the study were virtually identical across all conditions.

131 Mixon, T. R., & Moroney, W. F. (1982). An annotated bibliography of objective pilot performance measures (NAVTRAEQUIPCEN IH-330). Orlando, FL: Naval Training Equipment Center.

Buckout's review in 1962 was the last comprehensive examination of the pilot performance measurement (PPM) literature. This annotated bibliography attempts to: (1) gather the PPM literature written subsequent to 1962 into one source; (2) describe the scenarios and measures used in collecting PPM data; and (3) summarize the major premises and findings of each article. A variety of sources including computer aided literature search were used to identify candidate articles. Ultimately all referenced material was divided into three categories: (1) objective pilot performance measurement; (2) subjective pilot performance measures; and (3) general analysis and review articles. For each of the objective measure articles reviewed the following parameters were reported: subjects, equipment, scenario, measures and summary. For the subjective measures and general analysis and review articles the author's abstract was generally duplicated. In addition to the 189 articles addressing objective performance measurement, 30 articles dealing with subjective measures and 143 related analyses and review articles are contained in the bibliography. The authors hope that the document will (1) provide a means for integrating the PPM literature, and (2) serve as an impetus to develop a systematic approach to PPM.

132 Moray, N. (1980). Subjective measurement of mental workload (Tech. Rep. for Contract No. N00014-77-C-0256). Cambridge: Massachusetts Institute of Technology. (DTIC No. AD-A092 664)

Although there is widespread agreement that an important component of mental workload is the subjective judgement of how difficult the task seems, and how loaded the human operator feels, there have been rather few attempts to measure subjective load directly. This paper reviews those attempts and discusses what variables in the task seem to be responsible for the ratings given by the operator. Some suggestions for future research are given.

133 Moray, N. (1982). Subjective mental workload. Human Factors, 24, 25-40.

Subjective workload is of increasing importance in user-machine systems, as the role of the human operator becomes less to control and more to monitor complex systems. This paper reviews the relationship between physical, cognitive, manual-control, and time-stress tasks and the subjective load experienced by an operator. The more important task variables are suggested, as are areas requiring research.

134 Moray, N., Turksen, B., Aidie, P., Drascic, D., Eisen, P., Kruschelnicky, E., Money, L., Schonert, H., & Thornton, C. (1986). Progress in mental workload measurement. Proceedings of the Human Factors Society 30th Annual Meeting, 2, 1121-1123.

Two new techniques are described, one using subjective, the other physiological data for the measurement of workload in complex tasks. The subjective approach uses fuzzy measurement to analyze and predict the difficulty of combinations of skill based and rule based behaviour from the difficulty of skill based behaviour and rule based behaviour measured separately. The physiological technique offers an on-line real-time filter for measuring the Mulder signal at 0.1 Hz in the heart rate variability spectrum.

(From Proceedings of the Human Factors Society 30th Annual Meeting, 1986, 2, p. 1121. Copyright 1986 by the Human Factors Society, Inc. Reprinted by permission.)

135 Murphy, M. R., Randle, R. J., Tanner, T. A., Frankel, R. M., Goguen, J. A., & Linde, C. (1984). A full mission simulator study of aircrew performance: The measure of crew coordination and decisionmaking factors and their relationships to flight task performance. Proceedings of the 20th Annual Conference on Manual Control, NASA CP-2341, 2, 249-260.

Sixteen three-man crews flew a full-mission scenario in an airline flight simulator. The scenario was designed to elicit a high level of verbal interaction during instances of critical decisionmaking. Each crew flew the scenario only once, without prior knowledge of the scenario problem. Following a simulator run and in accord with formal instructions, each of the three crewmembers independently viewed and commented on a videotape of their performance. Two check-pilot observers rated pilot performance across all crews and, following each run, also commented on the video tape of that crew's performance. A linguistic analysis of voice transcripts is being made to provide added assessment of crew coordination and decisionmaking qualities. Measures of crew coordination and decisionmaking factors are being correlated with flight task performance measures. Some results and conclusions from observational data are presented.

136 National Aeronautics and Space Administration, Ames Research Center. (1986). NASA workload consultant for field evaluation (WC FIELDE). Moffett Field, CA: Author.

W.C. FIELDE is a microprocessor based system designed to assist users in selecting appropriate workload assessment procedures. It suggests measures, in descending order of utility, based on the users' answers to a variety of questions concerning their specific application. The factors that it takes into account include: the focus of the research question, the research environment, and the facilities that are available. It draws from a data base of widely used measures in proposing alternatives, and provides specific instructions about how

to apply many techniques. It was created with EXSYS, a commercially-available rule-based expert system development package. A copy protected version of the program is provided on the diskette. It runs on IBM/PC and IBM/PC compatible machines.

137 Nicholson, A. N. (Ed.). (1974). Simulation and study of high workload operations (AGARD-CP-146). Neuilly-Sur-Seine, France: North Atlantic Treaty Organization - Advisory Group for Aerospace Research and Development. (DTIC No. AD-A007 963)

The use of simulation for the evaluation of control dynamics, system components and procedures is well established; less certain is the use of such techniques in the evaluation of aircrew performance in high workload situations of an operational nature. Studies in the field present considerable difficulties, particularly when factors which may adversely affect flight safety have to be assessed. Although laboratory and airborne simulation of an operational situation may lack reality, such techniques may be the only way to predict operational performance. The meeting was intended to provide a series of presentations including analyses of operational aspects from several NATO countries.

(The original version of this material was first published by the Advisory Group for Aerospace Research and Development, North Atlantic Treaty Organization [AGARD/NATO] in CP 146, Simulation and Study of High Workload Operations, 1974. Reprinted by permission.)

138 Notestine, J. C. (1984). Subjective workload assessment and effect of delayed ratings in a probability monitoring task. Proceedings of the Human Factors Society 28th Annual Meeting, 2, 685-689.

In this investigation, twenty-four subjects performed a probability monitoring task under two difficulty levels and rated the workload associated with each using the Subjective Workload Assessment Technique (SWAT). Measures of response accuracy and response time in the monitoring task were also recorded. The effect of delaying SWAT ratings for both 15 and 30 minute intervals after task completion was also investigated. Study results indicated that SWAT was sensitive to the two levels of probability monitoring task difficulty, but not as sensitive as response accuracy. SWAT was, however, more sensitive to the varying levels than response time. Although there were no statistically significant effects associated with the interval between task completion and assigning ratings, five of the sixteen subjects who gave delayed ratings indicated that workload associated with the more difficult task was lower than the workload associated with the easier task. This finding may be interpreted as evidence that there may be practical consequences when ratings are delayed beyond task completion. The results, in general, are interpreted as supporting the applicability of SWAT as a sensitive workload measure.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 2, p. 685. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

139 O'Donnell, R. D. (1983). The U.S. Air Force neurophysiological workload test battery: Concept and validation. Sustained intensive air operations: Physiological and performance aspects (AGARD CP-338, pp. 5.1-5.9). Neuilly-Sur-Seine, France: North Atlantic Treaty Organization - Advisory Group for Aerospace Research and Development. (DTIC No. AD-A139 324)

In assessing the workload effects of sustained operations, it is likely that a multistage process will be employed. Broadly based measures such as timeline analyses and subjective estimates will be used to identify specific problem areas in a given sustained operation. These workload 'choke-points' will then be intensively studied utilizing a variety of subjective, behavioral, and physiological measures to tap the appropriate resources within the individual and to provide an estimate of how these resource demands interact with task demands and response variables. In order to provide measures of these specific capabilities and resources, the U.S. Air Force Aerospace Medical Research Laboratory is developing workload metrics spanning the entire spectrum, from task analysis through subjective and behavioral measures. In addition, neurophysiological measures have been investigated for a number of years with respect to their sensitivity in assessing workload. A test battery consisting of six different electrophysical measures in eleven different forms has been constructed. The tests selected were based on extensive laboratory data which indicated some probability that each measured an aspect of workload and might be appropriate and practical in applied settings. This test battery is currently undergoing validation studies in simulator environments and successfully validated tests will be incorporated into a second generation neurophysiological test battery to be used in field workload assessment.

(The original version of this material was first published by the Advisory Group for Aerospace Research and Development, North Atlantic Treaty Organization [AGARD/NATO] in CP 338, Sustained Intensive Air Operations: Physiological and Performance Aspects, 1983. Reprinted by permission.)

140 Parasuraman, R. (1985). Event-related brain potentials and intermodal divided attention. Proceedings of the Human Factors Society 29th Annual Meeting, 2, 971-975.

Attention allocation to visual and auditory channels under high-information load was examined by recording event-related brain potentials (ERPs). Ten subjects monitored an audiovisual display of intermittent 2-degree circles presented centrally and 1000-Hz tones presented binaurally. Subjects had to detect targets in both channels while dividing attention to each channel in varying proportions. Each subject had a minimum of 20 hours practice at the task. POC analysis indicated a tradeoff in processing resources between the visual and auditory channels. The N160 and P250 components of the visual ERP, and a slow negative shift potential associated with the auditory N100 component, varied in amplitude as processing resources were allocated to the visual or auditory channel. Both these sets of results were obtained only when stimuli were presented at a fast rate. The results suggest that intermodality divided attention influences both

modality-specific and modality-nonspecific ERP components in practised subjects under high-information load conditions. The implications of the results for models of processing resources and the evaluation of mental workload are discussed.

(From Proceedings of the Human Factors Society 29th Annual Meeting, 1985, 2, p. 971. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

141 Poock, G. K. (1972). Using pupillometry as an indicator of an operator's overload. Proceedings of the Human Factors Society 16th Annual Meeting, pp. 197-198.

Two experiments are described whose results tend to indicate that complexity and mental overload points of a job may be determined by analysis of pupillary responses.

(From Proceedings of the Human Factors Society 16th Annual Meeting, 1972, p. 197. Copyright 1972 by the Human Factors Society, Inc. Reprinted by permission.)

142 Rahimi, M., & Wierwille, W. W. (1982). Evaluation of the sensitivity and intrusion of workload estimation techniques in piloting tasks emphasizing mediational activity. Proceedings of the 1982 IEEE International Conference on Cybernetics and Society, pp. 593-597.

In this experiment, pilots flew an instrumented moving-base simulator. Mediational loading was elicited by having them solve a variety of navigational problems. The problems were sorted into low, medium, and high load conditions based on the number and complexity of arithmetic and geometric operations required to solve them. Workload estimation techniques based on opinion, spare mental capacity, primary task performance, and physiological measures were obtained and compared. This paper describes: (1) the ability of the techniques to discriminate statistically between the three levels of loading conditions, and (2) changes in primary task performance caused by introduction of the workload technique procedures and equipment.

(Copyright 1982 IEEE. Reprinted by permission from Proceedings of the 1982 IEEE International Conference on Cybernetics and Society, Seattle, WA, October 28-30, p. 593.)

143 Rehmann, J. T. (1982). Cockpit display of traffic information and the measure of pilot workload: An annotated bibliography (Report No. DOT/FAA/EM-81/9). Washington, DC: U.S. Department of Transportation, Federal Aviation Administration.

Approximately 80 references relating to pilot workload were selected and summarized as part of the Cockpit Display of Traffic Information (CDTI) studies currently being conducted by the Federal Aviation Administration technical Center in Atlantic City, New Jersey. A comprehensive search of the scientific literature was conducted using several sources, including books, scientific

journals, proceedings of technical meetings, and computerized information retrieval. Specific topics covered on this annotated bibliography, as they related to CDTI and its concomitant workload considerations, are subjective measures, spare mental capacity, primary task measures, and physiological measures.

144 Rehmann, J. T., Stein, E. S., & Rosenberg, B. L. (1983). Subjective pilot workload assessment. Human Factors, 25, 297-307.

Traditional subjective pilot workload measures have stressed postflight questionnaires. An alternative method that is less dependent on memory was evaluated in two experiments. In the first study, pilots and nonpilots made workload evaluations each minute during a critical tracking task. Results indicate that their responses were directly related to the experimentally controlled difficulty level, whereas posttask questionnaire responses were much less accurate. In a second study, the workload assessment device was introduced into the cockpit of a general aviation cockpit simulator, to determine if pilots could differentiate between three flights in which the level of difficulty varied. It was found that pilot workload judgments and response latencies were related to the experimentally induced difficulty level. As hypothesized, the more difficult flights generated higher mean workload responses and longer latencies.

(From Human Factors, 1983, 25, p. 297. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

145 Reid, G. B. (1985). Current status of the development of the subjective workload assessment technique. Proceedings of the Human Factors Society 29th Annual Meeting, 1, 220-223.

The Subjective Workload Assessment Technique (SWAT) has been under development for approximately five years. This measure is under a systematic development program to define its' strengths and weaknesses. Both laboratory research and field applications are being employed in this evaluation and some of the findings are presented. Current research on refinements to the procedure are discussed.

(From Proceedings of the Human Factors Society 29th Annual Meeting, 1985, 1, p. 220. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

146 Reid, G. B, Eggemeier, F. T., & Nygren, T. E. (1982). An individual differences approach to SWAT scale development. Proceedings of the Human Factors Society 26th Annual Meeting, pp. 639-642.

A refinement to the scale development phase of the Subjective Workload Assessment Technique (SWAT) provides for forming scales for homogeneous subject groups. Groups are formed by determining which of the three dimensions, time load, mental effort load or stress load subjects judge to be the most important contributor to workload. The group scales are then transformed into a SWAT scale that ranges

from 0 for the lowest defined workload condition to 100 for the highest workload condition. This procedure should increase the precision of workload measurement while minimizing the effects of individual subject ranking errors.

(From Proceedings of the Human Factors Society 26th Annual Meeting, 1982, p. 639. Copyright 1982 by the Human Factors Society, Inc. Reprinted by permission.)

147 Reid, G. B., Eggemeier, F. T., & Shingledecker, C. A. (1982). Subjective workload assessment technique. In M. L. Frazier & R. B. Crombie (Eds.), Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics (AFFTC-TR-82-5, pp. 281-288). Edwards Air Force Base, CA: Air Force Flight Test Center. (DTIC No. AD-A129 333)

Subjective assessment techniques are often employed in flight testing and operational test and evaluation. A thorough literature review of subjective measures revealed that precisely what we wanted does not currently exist. While subjective measures are frequently used in workload assessment, they usually are designed for a specific application. The measure we desire should be designed specifically to access workload. Within the context of our battery this measure is conceived as being less precise than some of the other proposed measures (e. g. cortical evoked response) but should be precise and sensitive enough to quantify the existence of high workload. To develop SWAT we defined workload as being primarily composed of three dimensions: Time load, Mental effort load, and Psychological stress load. Refinement of the technique will continue as we gain experience in a wide variety of applications in order to provide a general workload measure. Data is being collected regarding SWATs sensitivity, validity and reliability in order to aid potential users apply the technique to their own specific application.

148 Reid, G. B., Shingledecker, C. A., & Eggemeier, F. T. (1981). Application of conjoint measurement to workload scale development. Proceedings of the Human Factors Society 25th Annual Meeting, pp. 522-526.

Practical considerations make subjective opinion one of the most widely used methods to assess mental workload. However, the value of the data obtained by subjective methods is often limited because scales are not standardized and are not based on modern psychometric principles. This paper describes the development and validation of a Subjective Workload Assessment Technique (SWAT) which uses conjoint measurement to construct interval level workload scales from ordinal rankings of combinations of levels on three contributory dimensions. An experiment was conducted to investigate the construct validity and concurrent validity of the SWAT-1 scale. Recommendations for further development are discussed.

(From Proceedings of the Human Factors Society 25th Annual Meeting, 1981, p. 522. Copyright 1981 by the Human Factors Society, Inc. Reprinted by permission.)

149 Repperger, D. W., Rogers, D. B., Van Patten, R. E., & Frazier, J. (1982). A study of task difficulty with a subjective rating scale. In M. L. Frazier & R. B. Crombie (Eds.), Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics (AFFTC-TR-82-5, pp. 499-513). Edwards Air Force Base, CA: Air Force Flight Test Center. (DTIC No. AD-A129 333)

The results of two experiments are discussed which relate to taks difficulty and the effects of environmental stress on tracking performance. The first experiment involved 5 different sum of sine tracking tasks which humans tracked both in a static condition and under a 5 Gz acceleration stress condition. The tasks were designed in such a manner as to investigate workload measures and to compare our hypothetical design to subjective evaluations. The tasks were required to satisfy 5 criteria specified in mathematical terms. The second experiment involved similar environmental stress conditions but in this case the tasks were constructed from deterministic functions with specially designed velocity and acceleration profiles. In both parts of this experiment, subjective evaluations were obtained and compared to the assumption that difficulty is related to magnitudes of velocity and acceleration profiles of the target tracking task. Phase Plane performance analysis was conducted across 7 subjects to study potential measures of workload or tracking.

150 Revesman, M. E., & Rokicki, S. M. (1985). Assessment of SWAT accuracy. Proceedings of the Human Factors Society 29th Annual Meeting, 1, 183-187.

Recently, the Subjective Workload Assessment Technique (SWAT) has become a popular tool for assessing mental workload. SWAT attempts to estimate an individual's internal representation of mental workload from a rank ordering of combinations of mental workload factors. This paper presents an experiment in which these internal values were simulated on a computer. Conjoint analysis was performed on rank orders produced from the internal values with and without noise added. Results indicated that conjoint analysis was able to accurately estimate the "true" rankings and values, even when noise was present. In addition, linear regression was found to be a poor estimator of the "true" values, while a simpler technique, rank summing, was found to be as accurate an estimator as conjoint analysis.

(From Proceedings of the Human Factors Society 29th Annual Meeting, 1985, 1, p. 183. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

151 Robertson, M. M. (1984). Personality differences as a moderator of mental workload behavior: Mental workload performance and strain reactions as a function of cognitive complexity. Proceedings of the Human Factors Society 28th Annual Meeting, 2, 690-694.

Two groups were pre-selected on the basis of high abstraction or concreteness scores on a test for the personality trait of cognitive complexity. The groups of participants undertook four sequential phases of a fire command, computer-generated, mental workload task. Multiple measures of the mental workload were taken consisting of primary task performance, subjective estimate of relative task difficulty, and physiological response as reflected in the level of sinus

arrhythmia. Abstract individuals successfully completed the initial phase of the task, judged to be intermediate in workload, and the overall task in significantly fewer trials than did the concrete individuals. The abstract group also took significantly less time to complete the task during the initial acquisition phase, a trend that was affirmed in the overall time to complete the entire task. Concrete participants perceived the task to be most difficult during the initial phase, whereas the abstract group's subjective evaluation varied directly with the progressive loading of the task. Differences in the scores of sinus arrhythmia for the two groups were shown in the expected direction but failed to reach significance. This study indicates that personality, as reflected in the level of cognitive complexity, is an important moderator of human behavior in the performance of mental workload tasks.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 2, p. 690. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

152 Robertson, M. M., & Meshkati, N. (1985). Analysis of the effects of two individual differences classification models on experiencing mental workload of a computer generated task: A new perspective to job design and task analysis. Proceedings of the Human Factors Society 29th Annual Meeting, 1, 178-186.

Two individual differences classification models, Cognitive Complexity and Decision Style, were utilized to group participants for a mental workload experiment. All participants undertook four sequential phases of a simulated artillery command, computer generated, mental workload task. Measures of the mental workload were the subjective estimation of the relative task difficulty and the physiological response as reflected in the level of sinus arrhythmia. The abstract group was found to be significantly associated with the high conceptually complex decision styles; whereas the concrete group was related to the low conceptually complex decision styles. The abstract and conceptually complex decision style group subjective evaluation of task difficulty followed the progressive loading of the task. Conversely, the concrete and low conceptually complex decision style group perceived the initial phase of the task to be the most difficult. The sinus arrhythmia dependent variable demonstrated two different patterns for the unifocus and multifocus decisions styles. Implications of the results of this investigation could be employed by managers in developing a more effective job design approach and to optimize the job-person match. Further ramifications of the findings could contribute to efficient manpower planning and human resource utilization within the organization.

(From Proceedings of the Human Factors Society 29th Annual Meeting, 1985, 1, p. 178. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

153 Rosenberg, B., Rehmann, J., & Stein, E. (1982). The relationship between effort rating and performance in a critical tracking task (Report No. DOT/FAA/EM-81/13). Washington, DC: U.S. Department of Transportation, Federal Aviation Administration.

This report documents the results of a preliminary evaluation of a Pilot Objective/Subjective Workload Assessment Technique (POSWAT). The study employed a critical tracking task, in which 24 subjects (pilots and nonpilots) viewed an analog display of the error between operator input and system output while correcting with opposite pressure on a joystick. The purpose was to determine if there was a relationship between participant responses on a 10-point scale administered during task performance and tracking task difficulty. It is generally concluded that POSWAT used for measuring effort rating and rating delay on a regular basis during this experiment is minimally instrusive, is informative, and merits further evaluation in a cockpit environment.

154 Schiflett, S. G. (1980). Evaluation of pilot assessment device to test alternative display formats and control handling qualities (Report No. NATC-SY-33R-80). Patuxent River, MD: Naval Air Test Center.

This in-flight research project evaluated the utility of a workload assessment device (WAD) to measure pilot workload for approach and landing tasks under simulated instrument meteorological conditions, alternate HUD formats and control stability variations. The flight tests were conducted in a NT-33A research aircraft, extensively modified for the Air Force and Navy by the display evaluation flight test program. The hardware, software, and test procedures associated with the WAD functioned efficiently with only minor discrepancies and minimum pilot distraction. The project established the feasibility of using an item recognition task as a measure of sensory response loading and reserve information processing capacity while flying precision approaches. In a descriptive statistical treatment of the data, the results indicate appreciable increase in reaction time and errors with degraded handling qualities as compared to ground baseline measures and good handling qualities. The preliminary findings also reveal consistent trends toward the availability of more mental reserve capacity when flying predominantly pictoral/symbolic HUD configurations as compared to conventional HUD formats with scales and alphanumerics. It is recommended that further evaluations be conducted to establish the efficacy of utilizing the WAD to measure mental workload in a wide variety of aircrew tasks.

155 Schiflett, S. G. (1983). Theoretical development of an adaptive secondary task measure pilot workload for flight evaluations. Proceedings of the Human Factors Society 27th Annual Meeting, 1 602-605.

A series of research projects were reviewed that formulated the basis for the theoretical development and evaluation of an adaptive secondary task to measure pilot workload. The final flight project established the technical feasibility of using a visual and auditory item-recognition (Sternberg) task as a measure of sensory-response loading and reserve information processing capacity while flying precision pitch maneuvers simulating terrain profiles. The discrete item secondary task presented letters of the alphabet at a rate driven by a scoring algorithm that adapted to the pitch error scores i.e., cross-coupled to the

primary task. The preliminary results indicate an appreciable increase in reaction time and errors for the visual secondary task while flying the terrain avoidance primary task as compared to flying the same task under auditory task loading. Preliminary conclusions support the multiple resource model of information processing.

(From Proceedings of the Human Factors Society 27th Annual Meeting, 1983, 1, p. 602. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

156 Schiflett, S. G., & Loikith, G. J. (1980). Voice stress analysis as a measure of operator workload (TM 79-3 SY). Patuxent River, MD: Naval Air Test Center.

This study attempted to determine if the Psychological Stress Evaluator (PSE) could be used to detect the amount of situational stress in the voice while subjects performed a four-choice information processing task at different presentation rates. The 42 subjects were divided into group I - Jet, group II - Prop, and group III - Staff. A Response Analysis Tester (RATER) presented a four-choice discrimination task in which the subjects were required to match a response key to each of four stimuli (numbers - one, two, three, and four) appearing in a display window. The sequence of stimuli was randomly presented in an automatic-paced mode for nine 1-min tests. The stimuli presentation rates were set at one symbol per 1.5 sec, .75 sec, and .50 sec. During the first three tests, the subjects were instructed to press the correct key and not verbalize the number. During the next three tests, the subjects verbalized the number and did not press the key. During the last three tests, the subjects verbalized the number and simultaneously pressed the correct key. At the end of each block of three tests, the subjects estimated self- performance as percent correct and rated stress on a scale of one (no stress) to seven (high stress). Voice signals were initially recorded on magnetic tape, then processed through filtering circuits and displayed on a strip chart for subsequent visual analysis and interpretation. A subjective scoring criterion was established and then translated into electronic equivalents and automated on a Varian 73 computer for voice pattern recognition analysis. Significant main effects for percent-correct responses were obtained for groups, presentation rate, and groups X presentation rate interaction. No significant differences were found in the correct responses of the subjects when the number was verbalized or not verbalized. The Staff group produced significantly fewer correct responses than either the Jet or Prop groups at the .75 sec rate. Voice stress analysis showed significant correlations with performance scores and stress ratings of a selected pool of subjects (N=12). The results were discussed as to the potential application of an objective, reliable, sensitive, and nonobtrusive measure of stress in vocal communication systems that require operator workload assessments.

157 Schlegel, R. E., & Shingledecker, C. A. (1985). Training characteristics of the criterion task set workload assessment battery. Proceedings of the Human Factors Society 29th Annual Meeting, 2, 770-773.

An evaluation of the Criterion Task Set was performed to determine the training requirements for the various tasks. Twenty subjects were divided into four groups. One group trained on all nine tasks in the battery. The other three

groups trained on different three-task subsets. All subjects trained for two hours per day on five consecutive days. Response time, accuracy and subjective workload measures were obtained for each trial. The required number of trials for stable performance ranged from two to six with a mode of five. Slight improvements were observed on some tasks after eight to ten trials. Performance by the group trained on all nine tasks was equivalent on half of the tasks and worse on the other half. Subjective workload ratings were highly correlated with the actual performance scores.

(From Proceedings of the Human Factors Society 29th Annual Meeting, 1985, 2, p. 770. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

158 Senders, J. W. (1980). The problem of estimating workload (Report No. AFOSR-79-0133-#1). Santa Barbara: University of California.

The report presents a general theoretical discussion of some of the fundamental problems of analysis of human operator workload. Questions relating to the estimation as opposed to the measurement of workload are attacked. Any workload estimation technique as opposed to virtually all workload measurement techniques requires a comprehensive description of the human operator, the system to be used, and the mission of the man-machine system.

(Copyright 1980, University of California. All rights reserved. Reprinted by permission.)

159 Senders, J. W., & Gottsdanker, R. M. (1980). On the estimation of mental workload (Final Report for Contract No. AFOSR-79-0133). Santa Barbara: University of California, Department of Psychology. (DTIC No. AD-A094 783)

The only practical way to attack the problems of mission delimitation, machine improvement, and operator specification in respect to mental workload appears to be computer simulations of both man and system performing a well-specified mission. As real systems afford the operator relative freedom of choice, the simulation of the human operator must possess intelligence. Unfortunately the data gathered in the laboratory on human performance are highly synthetic and concern the lowest elements of behaviour. Uncritical use of these data to solve problems of mental workload cannot be justified. The utilization of methods of artificial intelligence to bear upon the intelligent restructuring of tasks appears feasible at the present time. It will probably suffice to include rational decision making and a capacity to solve certain classes of games.

(Copyright 1980, University of California. All right reserved. Reprinted by permission.)

160 Shingledecker, C. A. (1980). Enhancing operator acceptance and nonintereference in secondary task measures of workload. Proceedings of the Human Factors Society 24th Annual Meeting, pp. 674-677.

A serious disadvantage of many available secondary task measures of operator workload is that they are difficult to employ during later stages of system

design and evaluation. Common problems with traditional laboratory tasks include poor operator acceptance and the potential for intrusion on primary tasks. Two methods designed to overcome these limitations are being investigated at the Air Force Aerospace Medical Research Laboratory. In order to improve integration of the measurement task with the operator's duties, realistic aircraft radio communications activities have been developed as secondary tasks. Analytical and subjective methods have been employed to develop communications tasks with scaled workload values. A second project has focused on further development of the interval production task as a workload metric. This task has minimal response requirements and appears to act as an index of load rather than a reserve capacity test.

(From Proceedings of the Human Factors Society 24th Annual Meeting, 1980, p. 674. Copyright 1980 by the Human Factors Society, Inc. Reprinted by permission.)

161 Shingledecker, C. A. (1983). Behavioral and subjective workload metrics for operational environments. Sustained intensive air operations: Physiological and performance aspects (AGARD CP-338, pp. 6.1-6.8). Neuilly-Sur-Seine, France: North Atlantic Treaty Organization - Advisory Group for Aerospace Research and Development. (DTIC No. AD-A139 324)

The assessment of crew performance capability under conditions of sustained intensive air operations requires the use of specialized measures of operator workload which are matched to the nature of the investigation and to the environment in which the workload evaluation must be conducted. In many cases, the effects of severe combined stressors and of aircrew performance requirements on mental workload cannot be studied in the laboratory, and must be addressed in high fidelity simulation or during operational test exercises. This paper examines the advantages and limitations of traditional subjective report and behavioral measures of workload for application in operational environments. In addition, recent efforts at the U.S. Air Force Aerospace Medical Research Laboratory to develop improved field-usable subjective and behavioral secondary task metrics are described.

(The original version of this material was first published by the Advisory Group for Aerospace Research and Development, North Atlantic Treaty Organization [AGARD/NATO] in CP 338, Sustained Intensive Air Operations: Physiological and Performance Aspects, 1983. Reprinted by permission.)

162 Shingledecker, C. A. (1986). The criterion task set: Validation and application. Proceedings of the Human Factors Society 30th Annual Meeting, 1, 57 (abstract only).

The Criterion Task Set (CTS) is a battery of performance tasks which was developed at the Air Force Armstrong Aerospace Medical Research Laboratory. Based on an information processing stage/resource model of human performance, the CTS was designed to evaluate the relative sensitivity, diagnosticity and intrusiveness of available measures of operator workload. It has also been employed as a performance assessment instrument to evaluate the effects of stressors on hypothesized independent sources of performance capability. Since the completion of original developmental research and the implementation of the

CTS in a standard hardware/software system, a number of researchers have employed the battery in applied performance studies and in efforts which have contributed to its further refinement. The objectives of the symposium were to present accounts of six of their research projects and to provide a forum for individuals who are currently using the CTS or who are interested in potential applications of the performance assessment system.

(From Proceedings of the Human Factors Society 30th Annual Meeting, 1986, 1, p. 57. Copyright 1986 by the Human Factors Society, Inc. Reprinted by permission.)

163 Shingledecker, C. A., Crabtree, M. S., & Acton, W. H. (1982). Standardized tests for the evaluation and classification of workload metrics. Proceedings of the Human Factors Society 26th Annual Meeting, pp. 648-651.

Although a variety of metrics are available for workload assessment, little research has been conducted to construct guidelines for optimally matching them to individual workload questions or test environments. In an effort to remedy this problem, AFAMRL is currently developing a standardized methodology which will be used to evaluate candidate workload measures on several theoretical and pragmatic criteria. The central feature of this methodology is a set of representative loading tasks. This paper outlines the analyses that were conducted to screen potential components of the Criterion Task Set (CTS) and describes the tasks which were selected to produce demands on a number of functional dimensions of information processing.

(From Proceedings of the Human Factors Society 26th Annual Meeting, 1982, p. 648. Copyright 1982 by the Human Factors Society, Inc. Reprinted by permission.)

164 Shulman, H. G., & Greenberg, S. N. (1971). Perceptual deficit due to division of attention between memory and perception. Journal of Experimental Psychology, 88, 171-176.

Performance on a perceptual recognition task (Exp. 1) and a comparative judgment task (Exp. 2) was studied as a function of the information storage demands made by a concurrently performed short-term memory task. Performance on the perceptual tasks was inversely related to memory load both when recognition rate was the dependent variable and when reaction time was measured under conditions of nearly errorless performance.

165 Silverstein, L. D., Gomer, F. E., Crabtree, M. S., & Acton, W. H. (1984). A comparison of analytic and subjective techniques for estimating communications-related workload during commercial transport flight operations (NASA CR-2341). Washington, DC: National Aeronautics and Space Administration.

The objectives of this research contract were to develop a classification scheme for categorizing commercial transport communications and to apply analytic and subjective estimation techniques to quantify the workload imposed by these communications tasks. A communications task was defined as the sequence of

perceptual, cognitive, motor, and verbal responses initiated by the aircrew immediately following transmission of a message or instruction from ATC. Four techniques were used to quantify the workload: (1) an information-theoretic analysis, (2) paired-comparison technique for obtaining the opinions of current line pilots, (3) a combined hybrid scale that combined information from the other two techniques, and (4) a subjective rank-order scale. Highly significant agreement was found among the different methods of estimating communications workload. The results of this research provided a basis for the selection of standard sets of communications tasks with variable loading characteristics. Such a standard task repertoire can be used to control communications-related demands in future simulation research and should serve as input to a data-base of "workload calibrated" flight tasks.

166 Staveland, L., Hart, S. G., & Yeh, Y. Y. (1985). Memory and subjective workload assessment. Proceedings of the 21st Annual Conference on Manual Control, NASA CP-2428, pp. 7.1-7.13.

Recent research suggested subjective introspection of workload is not based upon specific retrieval of information from long-term memory, and only reflects the average workload that is imposed upon the human operator by a particular task. These findings are based upon global ratings of workload for the overall task, suggesting that subjective ratings are limited in ability to retrieve specific details of a task from long-term memory. To clarify the limits memory imposes on subjective workload assessment, the difficulty of task segments was varied and the workload of specified segments was retrospectively rated. The ratings were retrospectively collected on the manipulations of three levels of segment difficulty. Subjects were assigned to one of two memory groups. In the Before group, subjects knew before performing a block of trials which segment to rate. In the After group, subjects did not know which segment to rate until after performing the block of trials. The subjective ratings, RTs, and MTs were compared for within group, and between group differences. Performance measures and subjective evaluations of workload reflected the experimental manipulations. Subjects were sensitive to different difficulty levels, and recalled the average workload of task components. Cueing did not appear to help recall, and memory group differences possibly reflected variations in the groups of subjects, or an additional memory task.

167 Stein, E. S., Fabry, J., & Rosenberg, B. (1982). The elusive goal of measuring pilot workload in general aviation. In M. L. Frazier & R. B. Crombie (Eds.), Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics (AFFTC-TR-82-5, pp. 275-280). Edwards Air Force Base, CA: Air Force Flight Test Center. (DTIC No. AD-A129 333)

Techniques for measuring workload and performance in general aviation are under active development. The goal is to establish the tools necessary in order to provide timely and accurate information concerning the effects of systems changes on pilot behavior. Only through active, empirical research can such tools be developed.

168 Stern, J. A., & Skelly, J. J. (1984). The eye blink and workload considerations. Proceedings of the Human Factors Society 28th Annual Meeting, 2, 942-944.

Two parameters of the eye blink, blink rate and blink duration, were used to assess workload in two independent operational studies. Both studies involved high fidelity strategic bomber mission simulations. The first study was an extended wartime mission where workload was evaluated during mission segments. The second study involved shorter, discrete training missions where task difficulty was systematically manipulated. Both studies produced complementary results. Results show that: (1) blink rate is significantly affected by task demands; (2) blink rate is sensitive to task modality; (3) blink duration is significantly affected by task modality and complexity; and (4) blink duration is a sensitive index of time on task effects. These data support the use of eye blink measurement in "noisy" complex environments as both a feasible and valuable assessment technique.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 2, p. 942. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

169 Stone, L. W., & Duncan, C. E. (1984). Effects of extended use of AN/PVS-5 night vision goggles on helicopter pilots' performance (USARRL TR 84-3). Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory. (DTIC No. AD-A138 126)

The effects of extended use of AN/PVS-5 night vision goggles (NVG) were investigated by observing 10 NVG helicopter instructor pilots during two 6-hour missions. Each mission consisted of three 2-hour flights during which pilot control inputs and aircraft status variables were recorded in flight. Questionnaires were completed before the first mission and after the NVG mission. In order to examine for a carryover effect, subjects were flown in a crossover design in which half of the aviators flew NVG on the first mission, the other half on the second. Only the out-of-ground-effect hover showed a statistically significant carryover effect; that is, subjects who flew naked eye before NVG demonstrated a greater absolute difference in hover flight performance variability than those who flew naked eye after NVG flight. In the traffic pattern (final approach segment), there was a statistically significant difference between the visual conditions only. The postflight questionnaire responses revealed a concern over what was described as a "lack of concentration" and a "decline of mental alertness." Some physiological stress reactions were reported. None of the three maneuvers analyzed revealed a significant effect on performance across flights.

170 Sulzer, R. L., Cox, W. J., & Mohler, S. R. (1981). Flight crewmember workload evaluation (Report Nos. DOT/FAA/RD-82/21 & DOT/FAA/ASF-82/1). Washington, DC: U.S. Department of Transportation, Federal Aviation Administration. (DTIC No. AD-A114 167)

This is a report on transport category airplane flight crew workload measurement techniques as used in cockpit development and aircraft certification tests by major U.S. airframe manufacturers. It reviews the fundamentals of: crew size

certification; workload measures and criteria; workload studies made during aircraft design; and workload studies made after the design has been established, including those used in flight test. Certain documentation practices are identified. The limitations of the currently used practices and the needs for improved workload measurement techniques are addressed.

171 Tole, J. R., Stephens, A. T., Harris, R. L., & Ephrath, A. R. (1982). Quantification of pilot workload via instrument scan. In M. L. Frazier & R. B. Crombie (Eds.), Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics (AFFTC-TR-82-5, pp. 234-250). Edwards Air Force Base, CA: Air Force Flight Test Center. (DTIC No. AD-A129 333)

This paper describes work in progress on the use of visual scanning behavior as an indicator of pilot workload. The study is investigating the relationship between level of performance on a constant piloting task under simulated IFR conditions, the skill of the pilot, the level of mental workload induced by an additional verbal task imposed on the basic control task, and visual scanning behavior. The results indicate an increase in fixation dwell times, especially on the primary instrument with increased mental loading. Skilled subjects stared less under increased loading than did novice pilots. Sequences of instrument fixations were also examined. The percentage occurrence of the subject's most used sequences decreased with increased task difficulty for novice subjects but not for highly skilled subjects.

172 Ulrich, T. E., & Gorman, C. D. (1983). Job difficulty data as an indicator of job complexity. Proceedings of the 24th Annual Conference of the Military Testing Association, AFHRL-TP-83-16, pp. 665-670. (DTIC No. AD-A126 554)

It is widely acknowledged that the technological advances of the present are going to have a major impact on the design and operation of future weapons systems. Future systems will continue to depend on the performance of the operator of the maintenance person assigned responsibility for the system. Addressing this multifaceted problem calls for a focus on the attitudes of the people involved with respect to increasing complexity. It is clear that these attitudes can impact many elements of the system, from personnel manning to performance levels. The investigation of this area prompted the application of the Job Difficulty data obtained by the U.S.A.F. Occupational Measurement Center (USAFOMC) in a way which had not been used before.

173 Vidulich, M. A. (1986). Response modalities and time-sharing performance. Proceedings of the Human Factors Society 30th Annual Meeting, 1, 337-341.

The recent development of speech technology has provided an opportunity for new approaches in display/control design. Some researchers have proposed that the use of speech can reduce resource competition with manual controls and improve multi-task performance. However, it has also been suggested that due to the heavy reliance on within-subject experimental designs, the research supporting the resource competition hypothesis was potentially contaminated by asymmetric transfer. The present study examined the value of speech responses as a control

device in a dual-task experiment. The experimental design permitted the evaluation of asymmetric transfer effects. Despite numerous significant effects supporting the advantage of mixing manual and speech responses there was no statistically significant finding that suggested the occurrence of asymmetric transfer. Also, the value of speech output was demonstrated in between-subject analyses that were logically immune to asymmetric transfer effects. Therefore, although the possibility of asymmetric transfer remains a legitimate experimental design concern, it is not a sufficient explanation for the observed response modality effects. The present results supported the resource competition hypothesis of response modality effects, and suggested that in operational environments the judicious use of speech technology can enhance performance.

(From Proceedings of the Human Factors Society 30th Annual Meeting, 1986, 1, p. 337. Copyright 1986 by the Human Factors Society, Inc. Reprinted by permission.)

174 Vidulich, M. A., & Tsang, P. S. (1985). Assessing subjective workload assessment: A comparison of SWAT and the NASA-bipolar methods. Proceedings of the Human Factors Society 29th Annual Meeting, 1, 71-75.

Subjective assessments of workload are becoming increasingly important in the evaluation of new systems. Two popular methods were compared in the present investigation: (1) the Subjective Workload Assessment Technique (SWAT) which was developed around the use of conjoint analysis to create interval scales, and (2) a technique under development at NASA that utilizes an individually weighted workload score from a set of nine bipolar ratings. Both methods were applied in a laboratory experiment that required rating a number of single- and dual-task trials of tracking and/or a spatial transformation task. The dual transformation-tracking task results were reviewed. The results for the two assessment techniques were remarkably similar, indicating that the subjective experience of workload is sufficiently robust to be resistant to variations in the measuring technique. Also, both subjective assessment techniques were successful in measuring the differences in task difficulty as indicated by a multivariate analysis of performance. Finally, the specific strengths and weaknesses of each assessment technique were reviewed.

(From Proceedings of the Human Factors Society 29th Annual Meeting, 1985, 1, p. 71. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

175 Vidulich, M. A., & Wickens, C. D. (1982). The influence of S-C-R compatibility and resource competition on performance of threat evaluation and fault diagnosis. Proceedings of the Human Factors Society 26th Annual Meeting, pp. 223-226.

Stimulus/central processing/response compatibility defines the optimum assignment of tasks to input modalities (auditory, A and visual, V) and output modalities (manual, M and speech, S). Spatial tasks are S-C-R compatible with visual/manual assignments. Verbal tasks are compatible with auditory speech assignments. Ten subjects time-shared a spatial task of aerial threat evaluation with a verbal task of fault diagnosis. All four I/O modality combinations of the threat task were performed while the fault task was performed with A/M and V/M assignments.

The joint effects of compatibility, and competition between tasks for input and output modalities were demonstrated. When resource competition was held constant, the effects of compatibility were found to be enhanced in dual task conditions. When both influences varied they were demonstrated to counteract in certain conditions and balance each other's effect.

(From Proceedings of the Human Factors Society 26th Annual Meeting, 1982, p. 223. Copyright 1982 by the Human Factors Society, Inc. Reprinted by permission.)

176 Vidulich, M. A., & Wickens, C. D. (1983). Processing phenomena and the dissociation between subjective and objective workload measures (EPL-TR-83-2/ONR-TR-83-2). Urbana-Champaign: University of Illinois, Engineering-Psychology Research Laboratory.

Causes of dissociation between subjective workload assessments and objective performance were investigated. A Sternberg memory search task was utilized. Sternberg task configurations varied in the automaticity of performance, stimulus presentation rate, discernability of stimuli, and the value of good performance. Automaticity in Sternberg task performance was manipulated by using two independent sets of stimuli, one of which was consistently mapped (i.e. targets were always the same) while the other was inconsistently mapped (i.e. targets changed over trials). Also, all Sternberg configurations were performed both as single tasks and as part of dual-task combinations (with a manual control task). During testing subjects rated all trials on eight typical bipolar rating scales.

Analysis of the results detected three major differences (i.e. dissociations) between what the ratings of workload would predict and, the actual performance which occurred. Subjects' ratings: (1) did not reflect the dual-task advantage of the consistently mapped Sternberg, (2) predicted an advantage for the slower presentation rate in which performance was degraded, and (3) indicated a higher level of workload was associated with the performance gain in a bonus-available condition. All of these dissociations identified could potentially contaminate subjective assessments in the field. The results were interpreted as supporting cognitive-processing-based experimentation in subjective workload assessment aimed at identifying differences between the cognitive processing accounting for subjective assessments and those processes that produce performance.

177 Vidulich, M. A., & Wickens, C. D. (1984). Subjective workload assessment and voluntary control of effort in a tracking task. Proceedings of the 20th Annual Conference on Manual Control, NASA CP-2341, 2, 57-71.

A manual control tracking task was manipulated along two dimensions: (1) control order, and (2) forcing function bandwidth. In the first phase of the experiment subjective workload assessments were collected. It was found that subjective assessments of workload were closely associated with performance in the case of increasing control order, but not in the case of increasing bandwidth. This was interpreted as indicating that subjective workload assessments are most appropriate for the study of increasing difficulty centered in response selection processes as opposed to response execution processes. In the second phase of the experiment the subjects were asked to voluntarily limit the effort they applied in the performance of the tracking task. The results indicate that the subjects

were quite facile in doing this. However, comparison of this data to the findings of other studies that manipulated effort via dual-task biasing indicate that effort manipulation is much more potent in a single-task configuration. This finding is discussed in terms of multiple resource theories of attentional capacity. Also, the utility of an analysis of covariance (ANACOVA) procedure in studying the relationships between subjective ratings and performance is highlighted.

178 Vidulich, M. A., & Wickens, C. D. (1985). Causes of dissociation between subjective workload measures and performance: Caveats for the use of subjective assessments. Proceedings of the Third Symposium on Aviation Psychology, pp. 223-230.

Dissociations between subjective workload assessments and performance were investigated. The difficulty of a Sternberg memory search task was manipulated by varying stimulus presentation rate, stimulus discernibility, value of good performance, and automaticity of performance. All Sternberg task conditions were performed both alone and concurrently with a tracking task. Bipolar subjective workload assessments were collected. Dissociations between workload and performance were found related to automaticity, presentation rate, and motivation level. The results were interpreted as supporting the hypothesis that the specific cognitive processes responsible for subjective assessments can differ from those responsible for performance. The potential contamination these dissociations could inflict on operational workload assessments is discussed.

179 Warren, C. A., Stern, J. A., Eddy, D. R., Horst, R. L., Kramer, A. F., Parasuraman, R., Sanquist, T. F., & Wilson, G. F. (1985). The role of event-related potentials in human-machine applications. Proceedings of the Human Factors Society 29th Annual Meeting, 2, 981-985.

This paper has four aims: to provide a brief introduction to the topic; to indicate the consensus of the contributors concerning the nature and direction of the discussion; to characterize the contributors' attempts at understanding and forecasting some trends in event-related potential (ERP) research and applications, and, finally, to convey some degree of criticality regarding the papers to be presented.

(From Proceedings of the Human Factors Society 29th Annual Meeting, 1985, 2, p. 981. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

180 Way, T. C. (1981). Pilot workload - a practitioner's problem. Proceedings of the Human Factors Society 25th Annual Meeting, p. 53 (abstract only).

The selection of a new military or commercial aircraft is based, in part, on life-cycle cost or cost-of-ownership. These parameters include both acquisition and operating costs. In the latter, personnel cost is an increasingly important factor. This leads directly to reduction in crew size and automation or reallocation of tasks previously assigned to the now-missing crewmen. It is also common to find that a new aircraft is more complex than the one it replaced.

Frequently then, in new aircraft there are fewer crewmen and at least as many functions to be performed as in older aircraft. The challenge to designers of flight decks and avionics suites is to configure the crew station in such a way that required tasks can be accomplished with the allotted crew complement and workload is held within reasonable limits. The purpose of this paper is to outline the role of the human factors practitioner in meeting this challenge, particularly our role in predicting and measuring workload.

(From Proceedings of the Human Factors Society 25th Annual Meeting, 1981, p. 53. Copyright 1981 by the Human Factors Society, Inc. Reprinted by permission.)

181 White, S. A., MacKinnon, D. P., & Lyman, J. (1986). Modified Petri net model sensitivity to workload manipulations. Proceedings of the 21st Annual Conference on Manual Control, NASA CP-2428, pp. 3.1-3.17.

Petri nets have been used to represent systems with asynchronous, concurrent and parallel activities. These characteristics led some researchers to suggest the use of Petri nets in workload modeling where concurrent and parallel activities are common. Petri nets are represented by places and transitions. In the workload application, places represent operator activities and transitions represent events. Modified Petri Nets (MPNs) have been used to formally represent task events and activities of a human operator in a man-machine system. MPNs were used for example to model the tasks underlying the identification and reaction to a lube oil lead in a ship propulsion system, the checkout and start-up procedure for a Cessna 182 light aircraft. In addition it was used to formulate a MPN for POPCORN, a complex computer simulation at NASA-Ames for workload research. These descriptive applications demonstrate the usefulness of MPNs in the formal representation of systems. The purpose of this research is to investigate MPNs as a workload modeling tool, and it describes the results of an exploratory study of the sensitivity of MPNs to workload manipulations in a dual task.

182 Wickens, C. D. (1981). Processing resources in attention, dual task performance, and workload assessment (EPL-TR-81-3/ONR-TR-81-3). Urbana-Champaign: University of Illinois, Engineering-Psychology Research Laboratory. (DTIC No. AD-A102-719).

This report develops the concept of multiple resource theory in dual task performance and describes its relation to the measurement of operator work-load. Structural and capacity theories of attention and time-sharing are contrasted, and the latter are then elaborated to describe the quantitative relation between resources and performance, and the representation of dual task data by the performance operating characteristic within a resource framework. Some "inconsistencies" with a single resource (undifferentiated capacity) model of time-sharing are pointed out, and the multiple resources model is introduced. Data are presented supporting a specific model that defines resources by stages of processing, codes of processing, and modalities of encoding. This report discusses the relation between multiple resources and operator performance

strategies, and different measures of operator workload. The different implications of multiple resource theory on primary task, secondary task, and physiological and subjective measures of workload, and the relations between these are considered.

183 Wickens, C. D. (1984). The multiple resources model of human performance: Implications for display design. Human Factors Considerations in High Performance Aircraft (AGARD CP-371, pp. 17.1-17.6). Neuilly-Sur-Seine, France: North Atlantic Treaty Organization - Advisory Group for Aerospace Research and Development.

This paper describes three cognitive principles of display design proposed within the framework of the multiple resources model of human performance. These principles are: stimulus-central processing compatibility describing how the optimum association of display format (print, speech, graphics, sound, and lateral visual field) to the working memory code (spatial-verbal) used in performance of a task; resource competition, describing the optimum configuration for displaying two or more sources of task information is to employ separate resources; and task integration, a principle that constrains the application of the resource competition when separate stimulus elements must be integrated into a single mental model for the task. Five experiments are described that investigate these principles in isolation and in different combinations in aviation-related tasks. The data generally support all three principles, and indicate that compatibility is a dominant concept when placed in opposition with resource competition.

(The original version of this material was first published by the Advisory Group for Aerospace Research and Development, North Atlantic Treaty Organization [AGARD/NATO] in CP 371, Human Factors Considerations in High Performance Aircraft, 1984. Reprinted by permission.)

184 Wickens, C., Derrick, W., Berringer, D., & Micalizzi, J. (1980). The structure of processing resources: Implications for task configuration and workload. Proceedings of the Human Factors Society 24th Annual Meeting, pp. 253-256.

The processing resources enabling dual task performance are modeled as a vector quantity, with dimensions defined by processing stages, processing codes, and processing modalities. This multidimensional conception is employed to model operator workload and to predict task interference patterns employing a Sternberg task methodology. The implications to time-sharing performance are addressed as they pertain to the stage dimensions (Experiment 1: Manual Tracking), the code dimension (Experiment 2: Failure detection), and the code and modality dimension (Experiment 3: Multi-element monitoring).

(From Proceedings of the Human Factors Society 24th Annual Meeting, 1980, p. 253. Copyright 1980 by the Human Factors Society, Inc. Reprinted by permission.)

185 Wickens, C. D., & Kramer, A. (1985). Engineering psychology. Annual Review of Psychology, 36, 307-348.

Reviews research on the cognitive aspects of engineering psychology, issues in human-computer interaction, process control, and automation. Human performance limits are considered with respect to the following: perception; response processes; attention, including such factors as task configuration, individual differences in learning, and subjective and physiological measures of workload and decision making, including decision aids, heuristics and cognitive limits, and applications to criminal justice and forecasting. Other issues considered include causal inference and diagnosis and errors and internal models in human performance. Process control is discussed in terms of the nature of alarm indicators, the diagnostic process, and diagnosis training. Human-computer interaction is examined in relation to the psychology of programming, the learning of text editors, data manipulation and retrieval, and the implementation of automation.

186 Wickens, C. D., Mountford, S. J., & Schreiner, W. (1981). Multiple resources, task-hemispheric integrity, and individual differences in time-sharing. Human Factors, 23, 211-229.

Forty subjects performed four information processing tasks. Three aspects of the data were examined: individual differences in time-sharing efficiency; task-related differences in time-sharing efficiency; and interaction of processing codes and response hands. Results showed little concrete evidence for a general time-sharing ability; supported a multiple resource concept of attention; and favored a condition of "task-hemispheric integrity" in which the hemisphere of task processing (spatial or verbal) is the same one controlling the task responding hand (left vs right). The implications of these results to operator performance in complex systems are discussed.

187 Wickens, C. D., Sandry, D. L., & Vidulich, M. A. (1983). Compatibility and resource competition between modalities of input, central processing, and output. Human Factors, 25, 227-248.

Synthesized auditory displays and speech recognizers were used in two experiments to develop guidelines for their implementation in military aircraft. In the first experiment, the competition between encoding and response modalities of concurrent tasks was examined. The memory search task was more susceptible to competition for visual encoding, whereas, the tracking task bore the greater impact from shared manual responding. The second experiment examined competition between tasks for encoding and response modalities and the optimum assignment of modalities to a given task. A simulated flight task was performed concurrently with either a spatial task (target acquisition) or a verbal task (memory). Best

performance and least interference with the flight task were obtained when the spatial task was displayed visually and responded to manually and also when the verbal task was displayed auditorily and responded to with speech.

(From Human Factors, 1983, 25, p. 227. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

188 Wickens, C. D., Vidulich, M. A., & Sandry-Garza, D. (1983). Principles of S-C-R compatibility with spatial and verbal tasks. Proceedings of the Second Symposium on Aviation Psychology, pp. 299-306.

A pilot's tasks may be categorized into those that demand predominantly verbal operations and those that are spatial. We describe two experiments that define two principles of compatibility of interfacing such tasks with displays and controls. The first defines compatibility according to display-location and response hand; the second according to the modality of display (auditory and visual) and response (manual and speech). In both experiments, these principles of compatibility are confirmed under dual task conditions. We describe their implications for cockpit design.

189 Wickens, C. D., & Yeh, Y. Y. (1983). The dissociation between subjective workload and performance: A multiple resource approach. Proceedings of the Human Factors Society 27th Annual Meeting, 1, 244-248.

A theory of the dissociation between subjective measures of mental workload and performance is described. The theory proposes that subjective measures are heavily driven by the number of tasks or task elements that a subject must perform concurrently. However they are relatively less sensitive to whether these tasks compete for common or separate resources, and to the difficulty of a single task, particularly if this difficulty is related to response factors. Performance, on the other hand, is particularly influenced by single task difficulty of both a perceptual and response nature and by resource competition between tasks. A set of three experiments are described to examine the dissociation between subjective difficulty measures and performance. These experiments employ different combinations of three tasks: Tracking, memory search, and a simulated air traffic control problem. The results supported all forms of dissociation predicted by the theory and the implications of results to workload measurement are discussed.

(From Proceedings of the Human Factors Society 27th Annual Meeting, 1983, 1, p. 244. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

190 Wickens, C. D., & Yeh, Y. Y. (1986). A multiple resources model of workload prediction and assessment. Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, 2, 1044-1048.

The Multiple Resource Model defines different structural resources within the human processing system. This paper first describes how the model may be employed, early in the system design process, to predict performance in complex settings. Limitations of the model in this regard are also pointed out. The

paper then describes how the model may be used for prescribing workload assessment techniques late in the design process, and for interpreting the dissociations that are often observed between subjective workload and performance.

(Copyright 1986 IEEE. Reprinted by permission from Proceedings of the 1986 IEEE International Conference on Systems, Man, and Cybernetics, Atlanta, GA, Oct. 11-13, p. 1044.)

191 Wiener, E. L., Curry, R. E., & Faustina, M. L. (1984). Vigilance and task load: In search of the inverted U. Human Factors, 26, 215-222.

The "Inverted-U Hypothesis" states that for a given task, there is an optimal level of workload or demand that yields the highest level of performance. A departure in either direction will result in a monotonically lower performance level, hence an inverted-U-shaped relationship between task demand and quality of performance. Most studies to date have failed to demonstrate the left-hand branch of the curve, that is, the regime in which performance presumably rises as load increases. The purpose of this study was to explore whether low-level additional demand on the monitor would result in improved performance. Four groups of subjects performed a visual monitoring task for 48 min, then two of the four groups were given additional tasks, and a third had potentially distracting information on its display. Results indicated that the two groups with additional demand detected more signals than did the control group or the control-plus-distraction group. There were no significant differences in false alarms.

(From Human Factors, 1984, 26, p. 215. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

192 Wierwille, W. W. (1982). Determination of sensitive measures of pilot workload as a function of the type of piloting task. In M. L. Frazier & R. B. Crombie (Eds.), Proceedings of the Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics (AFFTC-TR-82-5, pp. 471-490). Edwards Air Force Base, CA: Air Force Flight Test Center. (DTIC AD-A129-333)

The purpose of our present work, sponsored by NASA-AMFS, is to examine the sensitivity, intrusion, and transferability of a variety of workload assessment techniques. The study will use four different simulated piloting tasks, emphasizing psychomotor, perceptual, mediational, and communications aspects. Pilot loading levels will be systematically adjusted. Our simulation facility is a GAT-1B that has been modified and instrumented for workload estimation techniques measurement. The flight simulator itself has three degrees of physical motion and a full complement of IFR instruments. Recently we completed the experiment emphasizing the psychomotor aspect of flight. Instrument-rated pilots flew instrument approaches under three combined settings of the independent variable: increasing turbulence and decreasing longitudinal stability. Twenty different workload measures were taken between the outer and middle markers, only five of which showed statistically reliable changes as a function of the independent variable. Included in the five were: two rating scales, one measure of control movement activity, pulse rate, and one measure of time estimation. The results of the experiment are to some extent surprising,

for they indicate that several "accepted" measures of workload are not reliably sensitive to the kinds of psychomotor load which pilots encounter.

193 Wierwille, W. W. (1983). Comparative evaluation of workload estimation techniques in piloting tasks (NASA CR-166496). Washington, DC: National Aeronautics and Space Administration.

The objective of this research was to examine the sensitivity and intrusion of a wide variety of workload assessment techniques in simulated piloting tasks. The study employed four different piloting tasks emphasizing psychomotor, perceptual, mediational, and communications aspects of piloting behaviors. An instrumented moving base general aviation aircraft simulator was used for the study. This document provides a summary of the research.

194 Wierwille, W. W., & Casali, J. G. (1983). A validated rating scale for global mental workload measurement applications. Proceedings of the Human Factors Society 27th Annual Meeting, 1, 129-133.

The Cooper-Harper (1969) scale has been extensively used for evaluation of aircraft handling qualities and associated mental workload. The scale is a 10-point scale with a decision tree. A modified version of the scale, called the MCH scale, has been devised for the purpose of assessing workload in systems other than those where the human operator performs motor tasks; namely, where perceptual, mediational, and communications activity is present. The MCH scale has been validated in three different experiments. The scale is recommended for applications in which overall mental workload is to be assessed.

(From Proceedings of the Human Factors Society 27th Annual Meeting, 1983, 1, p. 129. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

195 Wierwille, W. W., & Connor, S. A. (1983). Evaluation of 20 workload measures using a psychomotor task in a moving-base aircraft simulator. Human Factors, 25, 1-16.

The sensitivity and intrusion of 20 pilot workload assessment techniques were compared using a psychomotor loading task in a three degree-of-freedom moving-base aircraft simulator. The primary task was an instrument landing system approach and landing, with measures taken between the outer and middle markers. Three levels of psychomotor load were obtained by combined manipulation of random wind-gust disturbance level and pitch stability. Two rating scale measures and one control movement measure demonstrated sensitivity to all levels of load. Additionally, one time estimation measure and one pulse rate measure demonstrated sensitivity to some levels of load. No intrusion was found. The results of this experiment indicate that the sensitivities of workload estimation techniques vary widely, and that only a few techniques appear sensitive to psychomotor load.

(From Human Factors, 1983, 25, p. 1. Copyright 1983 by the Human Factors Society, Inc. Reprinted by permission.)

196 Wierwille, W. W., Rahimi, M., & Casali, J. G. (1985). Evaluation of 16 measures of mental workload using a simulated flight task emphasizing mediational activity. Human Factors, 27, 489-502.

As aircraft and other systems become more automated, a shift is occurring in human operator participation in these systems. This shift is away from manual control and toward activities that tap the higher mental functioning of human operators. Therefore, an experiment was performed in a moving-base flight simulator to assess mediational (cognitive) workload measurement. Specifically, 16 workload estimation techniques were evaluated as to their sensitivity and intrusion in a flight task emphasizing mediational behavior. Task loading, using navigation problems presented on a display, was treated as an independent variable, and workload-measure values were treated as dependent variables. Results indicate that two mediational task measures, two rating scale measures, time estimation, and two eye behavior measures were reliably sensitive to mediational loading. The time estimation measure did, however, intrude on mediational task performance. Several of the remaining measures were completely insensitive to mediational load.

(From Human Factors, 1985, 27, p. 489. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

197 Wierwille, W. W., Skipper, J. H., & Rieger, C. A. (1984). Decision tree rating scales for workload estimation: Theme and variations. Proceedings of the 20th Annual Conference on Manual Control, NASA CP-2341, 2, 73-84.

The Modified Cooper-Harper (MCH) scale has been shown to be a sensitive indicator of workload in several different types of aircrew tasks (Wierwille and Casali. 1983). The study to be described in this paper was undertaken to determine if certain variations of the scale might provide even greater sensitivity and to determine the reasons for the sensitivity of the scale. The MCH scale, which is a 10 point scale, and five newly devised scales were examined in two different aircraft simulator experiments in which pilot loading was treated as an independent variable. The five scales included a 15 point scale, computerized versions of the MCH and 15 point scales, a scale in which the decision tree was removed, and one in which a 15 point left-to-right format was used. The results of the study indicate that while one of the new scales may be more sensitive in a given experiment, task dependency is a problem. The MCH scale on the other hand exhibits consistent sensitivity and remains the scale recommended for general use. The MCH scale results are consistent with earlier experiments also. This paper presents the results of the rating scale experiments and also describes the questionnaire results which were directed at obtaining a better understanding of the reasons for the relative sensitivity of the MCH scale and its variations.

198 Wierwille, W. W., & Williges, B. H. (1980). An annotated bibliography on operator mental workload assessment (NATC-SY-27R-80). Patuxent River, MD: Naval Air Test Center.

An annotated bibliography on operator mental workload is presented with supporting information. This bibliography is based upon two literature searches, one performed in 1977 in support of a survey and analysis catalog (AD A059-501)

and one performed in 1979 as an update. Each literature citation presented contains reference information, an abstract, a numerical workload technique category classification, a numerical operator behavior classification, and a group of word descriptors. Workload methods are divided into 28 specific techniques in four major categories: opinion, spare mental capacity, primary task, and physiological. Applicable operator behaviors are similarly divided into categories. The descriptors associated with each citation designate the general workload classification, the specific workload classification, the type of presentation, the type of facilities used, and the potential aircrew application. Over 600 citations are presented. Two indexes are also provided. The first is a workload technique index and the second is an experimental facility index. It is concluded that periodic updating of the bibliography will be required and that attention should be directed toward computerizing future workload bibliographies.

199 Wierwille, W. W., & Williges, R. C. (1978). Survey and analysis of operator workload assessment techniques (NATC-TR-SY-78-101). Patuxent River, MD: Naval Air Test Center. (DTIC No. AD-A059 501)

Over 400 references relating to operator mental workload were selected and classified according to a two-dimensional scheme including workload methodology and universal operator behavior. Twenty-eight specific techniques of assessing workload by means of subjective opinions, spare mental capacity, primary task, and physiological measures were cataloged. This catalog summarizes critical criteria that need to be considered in the flight test and evaluation environment and describes each technique in terms of theory and background, description of necessary method/apparatus, area of application and example, limitations, and suggested RDT & E follow-ups.

200 Williams, L. J. (1982). Cognitive load and the functional field of view. Human Factors, 24, 683-692.

In an experiment that kept visual display factors constant but which varied cognitive load, it was found that cognitive load modulated the functional field of view. When given a high level of foveal (cognitive) load, the functional field of view was only about 2 deg in diameter, whereas a low level of foveal load resulted in a functional field of about 4 deg diameter. The shrinkage of the functional field appeared to be rather generalized and not a true tunnel-vision effect.

(From Human Factors, 1982, 24, p. 683. Copyright 1982 by the Human Factors Society, Inc. Reprinted by permission.)

201 Williams, L. J. (1985). Tunnel vision induced by a foveal load manipulation. Human Factors, 27, 221-227.

In a tachistoscopic dual-task study, the foveal load of a primary task was varied, and subjects were required to report the orientation of a simultaneously presented peripheral line that could occur at any one of six retinal locations (3, 6, or 9 deg from central fixation in either the left or right visual field). Overall peripheral (secondary) task accuracy was significantly superior for

subjects given a low foveal load that required naming a letter drawn from a set of two alternatives as compared with the accuracy of subjects who had to choose a foveal response from among six alternatives. Foveal load interacted strongly with retinal eccentricity. Accuracy of peripheral line-orientation discrimination deteriorated much more rapidly with increases in retinal eccentricity for the high-foveal-load group than for the low-foveal-load group. These results are discussed in terms of a tunnel-vision model.

(From Human Factors, 1985, 27, p. 221. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

202 Wilson, G. F. (1985). A neuropsychological test battery for workload assessment. Proceedings of the Human Factors Society 29th Annual Meeting, 1, 224-225.

A Neuropsychological Workload Test Battery has been developed at AAMRL. It consists of eleven tests which were chosen to be appropriate for testing human responses in a wide range of applied situations. The tests incorporated into the battery measure averaged brain wave responses, heart, eye and muscle activity. The first version of the battery is currently undergoing testing in simulator environments.

(From Proceedings of the Human Factors Society 29th Annual Meeting, 1985, 1, p. 224. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

203 Wortman, D. B., Seifert, D. J, & Duket, S. D. (1975). SAINT simulation of a remotely piloted vehicle/drone control facility (AMRL-TR-75-119). Wright-Patterson Air Force Base, OH: Aeromedical Research Laboratory. (DTIC No. AD-A029 9446)

SAINT is a modeling and simulation technique that provides the concepts necessary to model systems that contain tasks (discrete elements), state variables (continuous elements), and interactions between them. SAINT has been designed to facilitate the modeling and analysis of complex man-machine systems. This paper describes a SAINT network model of a real-time simulation of a drone control facility (DCF) in which operators monitor and control the flight of simulated remotely piloted vehicles (RPVs) through the use of visual (CRT) displays.

204 Yeh, Y. Y., & Wickens, C. D. (1984). Why do performance and subjective workload measures dissociate? Proceedings of the Human Factors Society 28th Annual Meeting, 1, 504-508.

A set of three experiments are described that examine the sources of information processing that produce a dissociation between subjective workload measures and performance. The results support a theory of the dissociation. Subjective measures are driven more by the number of tasks currently performed. Subjective measures are also less sensitive to resource competition than performance

measures are. Factors that demand more resource investment improve performance, but these factors also increase subjective ratings of workload.

(From Proceedings of the Human Factors Society 28th Annual Meeting, 1984, 1, p. 504. Copyright 1984 by the Human Factors Society, Inc. Reprinted by permission.)

205 Yeh, Y. Y., & Wickens, C. D. (1985). An investigation of the dissociation between subjective measures of mental workload and performance (EPL-TR-84-1/NASA-TR-84-1). Urbana-Champaign: University of Illinios, Engineering-Psychology Research Laboratory.

The report addresses the dissociation between subjective measures of mental workload and performance. Three generic factors are identified that will drive subjective workload upward more than driving performance downward: Perceptual (versus response) load, and increased number of tasks, and better data quality. One factor, resource competition, is assumed to drive performance more than subjective workload. The theory of dissociation is tested in three experiments that employ different variations and combinations of three different tasks (tracking, memory search, and a simulated air traffic control task). The predictions of the theory are generally supported by the data. In addition, various subjective scales of mental workload are tested across the experiments. The correlations between these scales and multidimensional scaling data are used to help interpret the hidden cognitive structure of task difficulty.

206 Yeh, Y. Y., Wickens, C. D., & Hart, S. G. (1985). The effect of varying task difficulty on subjective workload. Proceedings of the Human Factors Society 29th Annual Meeting, 2, 765-769.

The goal of the present study was to determine whether or not retrospective workload ratings would reflect the average demands of the entire block of trials or whether one segment within the block would have more weight in determining the magnitudes of ratings than another. Performance data within a block of trials almost perfectly reflected the different task difficulty manipulations: reaction times (but not movement times) reflected variations in the difficulty of the more cognitive resopnse selection component whereas movement times (but not reaction times) reflected variations in the difficulty of the response execution component. Subjective ratings consistently reflected the combined demands of both task components averaged across levels of difficulty even when their levels of difficulty were varied within the block of trials. In every case, it appeared that all of the trials within a block were given equal weight in the composite subjective evaluation. These results suggest that subjective workload is not a specific retrieval of experiences heeded in working memory. Rather, it may reflect the experiences of an ongoing integration process.

(From Proceedings of the Human Factors Society 29th Annual Meeting, 1985, 2, p. 765. Copyright 1985 by the Human Factors Society, Inc. Reprinted by permission.)

AUTHOR INDEX

A

B

C

D

E

F

G

H

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

Y

Z

SUBJECT INDEX

GENERAL

MODELING

PERFORMANCE

Error

Mediational

Primary

Secondary

PHYSIOLOGICAL

006, 011, 016, 025, 026, 032, 054, 059, 067, 071, 078, 082, 086, 087, 093, 098, 099, 108, 111, 119, 120, 122, 123, 127, 134, 139, 140, 141, 142, 143, 147, 151, 152, 168, 169, 182, 185, 196, 198, 199, 202

Evoked Potential/P300

016, 067, 071, 119, 120, 140, 147, 179

Eye Movement/Eye Blinks/Pupil Size

011, 025, 098, 141, 168, 169, 196, 202

Heart Rate/Blood Pressure

078, 087, 093, 099, 122, 123, 127, 134, 202

Respiration Rate

006, 026, 078

Skin Resistance

078, 122, 123

SUBJECTIVE

005, 006, 007, 009, 010, 012, 013, 014, 015, 016, 017, 018, 019, 025, 026, 030, 031, 032, 036, 038, 039, 040, 041, 042, 044, 045, 046, 049, 050, 053, 054, 055, 056, 057, 058, 059, 060, 062, 066, 071, 072, 073, 074, 075, 078, 083, 086, 087, 088, 089, 090, 091, 092, 093, 094, 099, 100, 101, 103, 107, 111, 114, 115, 124, 125, 129, 130, 132, 133, 134, 138, 139, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 156, 157, 160, 161, 165, 166, 174, 176, 177, 178, 182, 185, 189, 190, 192, 194, 195, 196, 197, 198, 199, 204, 205, 206

Bipolar

031, 045, 046, 094, 099, 101, 129, 130, 174, 176, 178

Cooper-Harper

025, 026, 040, 103, 194, 197

Interviews/Opinions

006, 007, 012, 026, 036, 066, 089, 094, 142, 148, 165, 198, 199

Rating Scales

010, 013, 014, 015, 016, 017, 018, 030, 031, 041, 050, 072, 074, 078, 086, 087, 088, 089, 091, 092, 093, 094, 099, 100, 101, 107, 111, 114, 115, 132, 149, 156, 157, 166, 176, 177, 192, 194, 195, 196, 197, 204, 206

Subjective Workload Assessment Technique (SWAT)

007, 019, 038, 039, 056, 057, 058, 059, 060, 138, 145, 146, 147, 148, 150, 153, 174

REFERENCES

Wierwille, W. W., & Williges, B. H. (1980). An annotated bibliography on operator mental workload assessment (NATC-TR-27R-80). Patuxent River, MD: Naval Air Test Center.