

DAA/AMES

ENGINEERING-PSYCHOLOGY RESEARCH LABORATORY

University of Illinois at Urbana-Champaign

TECHNICAL REPORT EPL-84-2/NASA-84-2

OCTOBER 1984

The Dissociation of  
Subjective Measures of  
Mental Workload and  
Performance

FINAL REPORT

Yei-Yu Yeh  
Christopher D. Wickens



Prepared for

National Aeronautics and Space Administration  
NASA Ames Research Center  
Moffett Field, CA  
Contract No. NASA-NA-G-2-169

Approved for Public Release: Distribution Unlimited

(NASA-CR-176609) THE DISSOCIATION OF  
SUBJECTIVE MEASURES OF MENTAL WORKLOAD AND  
PERFORMANCE Final Report (Illinois Univ.,  
Urbana-Champaign.) 98 p HC A05/MF A01  
N86-21139 Unclas 15127  
CSCI 051 G3/53

## Table of Contents

	page
Abstract . . . . .	iii
Introduction . . . . .	1
The multiple resources model . . . . .	2
Subjective Workload . . . . .	4
The structure of perceived workload: A Multidimensional Scaling approach . . . . .	5
A model of verbal report data . . . . .	6
A theory of dissociation between performance and subjective workload . . . . .	8
Experimental Overview . . . . .	12
Hypotheses of the dissociation . . . . .	13
Secondary issues . . . . .	17
Methodology of the analysis . . . . .	19
Test of dissociation . . . . .	20
Workload measures in comparable units . . . . .	21
Test of the resource demand of a task: The additive factors methodology . . . . .	23
Experiment 1 . . . . .	26
Method . . . . .	27
Subjects . . . . .	27
Apparatus . . . . .	27
Tasks . . . . .	27
Procedure . . . . .	29
Experiment 2 . . . . .	30
Method . . . . .	31
Subjects . . . . .	31
Tasks . . . . .	31
Procedure . . . . .	33
Experiment 3 . . . . .	34
Method . . . . .	35
Subjects . . . . .	35
Tasks . . . . .	35
Procedure . . . . .	36
Results . . . . .	37
The processing demands of various tracking tasks . . . . .	37
Experiment 1 . . . . .	37
Experiment 2 . . . . .	41
Experiment 3 . . . . .	44
Rating reliability . . . . .	46
Predicting global workload from specific rating scales . . . . .	47
The perceived structure of subjective difficulty . . . . .	48
Dissociation between performance and subjective measures . . . . .	51
Perceptual vs. Response resources . . . . .	52
The number of display elements . . . . .	54
Single task difficulty vs. dual task competition . . . . .	56
Dual task competition for common vs. separate resources . . . . .	58
The effect of resource-investment factors . . . . .	61

	page
Discussion . . . . .	64
MDS results . . . . .	67
Discussion of dissociations . . . . .	69
Strength of dissociations . . . . .	77
Subjective workload scales . . . . .	79
Conclusion . . . . .	84

## Abstract

A dissociation between performance and subjective workload measures occurs, when two task configurations are compared and one shows better performance, but is perceived as subjectively more difficult than the other. The dissociation phenomenon was investigated in the theoretical framework of the multiple resources model. Even though the underlying structure of subjective workload strongly corresponds with the structure of processing resources, subjective measures do not preserve the vector characteristics in the multidimensional space described by the model. A theory of dissociation (Wickens & Yeh, 1983) was proposed to locate the sources that may produce dissociation between the two workload measures. According to the theory, performance is effected by every aspect of processing whereas subjective workload is sensitive to the amount of aggregate resource investment and is dominated by the demands on the perceptual/central resources. The proposed theory was tested in three experiments, employing different combinations of a tracking task and a Sternberg memory search task.

In supporting the theory, the results showed that performance improved but subjective workload was elevated with an increasing amount of resource investment. Furthermore, subjective workload, being affected by the aggregate demands, was not as sensitive as was performance to differences in the amount of resource competition between two tasks. The demand on perceptual/central resources was found to be the most salient component of subjective workload from both the multidimensional analysis of the hidden structure and the regression analysis of the underlying components. Dissociation occurred when the demand on this component was increased by the number of concurrent tasks or by the number of display elements. However,

in contrast to the prediction, demands on response resources were weighted in subjective introspection as much as demands on perceptual/central resources. The implications of these results for workload practitioners are described.

## Introduction

The concept of mental workload is of paramount importance in designing man-machine systems. The main concern of a designer is to enhance performance by moderating the workload experienced by the operators. Numerous techniques have been proposed to study and define the concept of mental workload. Wierwille and Williges (1978) classified these techniques into three categories: performance, physiological, and subjective assessment.

Despite the fact that all of these techniques are reputed to measure the same hypothetical construct, they provide different workload values in many conditions. Given the same manipulations, one technique may indicate an increase whereas another technique may indicate no change or a decrease in workload (e.g., Hicks & Wierwille, 1979; Williges & Connor, 1983). Such dissociations lead to various operational definitions of mental workload.

Even though there is no consensus with regard to the definition of mental workload, it is agreed that workload is multidimensional. Mental workload can only be specified as a vector in a to-be-defined space (Johannsen, Moray, Pew, Rasmussen, Sanders, & Wickens, 1979). To define the multidimensional space of mental workload, the underlying dimensions must be identified. Two models have presented candidate dimensions to define such a space.

The model proposed by Sheridan and Simpson (1979) has been accepted by many researchers as a framework for assessing subjective workload. According to this model, workload is defined by three descriptive categories: (1) task time constraints, (2) task uncertainty and complexity

of planning, and (3) psychological stress. However, these three dimensions were originally chosen through an intuitive subjective task analysis and the theoretical basis was not provided. Although these scales have undergone further development and refinement (Reid, Schingledacker & Eggemeier, 1981), it is still not entirely clear how these dimensions relate to human information processing and performance. It is also uncertain whether each dimension is itself multidimensional and how the three dimensions relate to each other (Boyd, 1982).

#### The multiple resources model

As an alternative approach, the dimensions of mental workload may be defined by the structure that characterizes human information processing and performance. According to this approach, workload is multidimensional because the human information processing system is multichanneled (Sanders, 1979) and depends upon separate resources of limited quantity (Navon & Gopher, 1979). Wickens (1981, 1984) identified three dichotomous dimensions as the functional composition of this processing system. These dimensions are (1) stages of processing (perceptual/central vs. response selection and execution), (2) modality of input (visual vs. auditory) and of output (manual vs. vocal speech), and (3) codes of central processing (spatial vs. verbal).

A hypothetical structure of these dimensions is shown in Figure 1. However, these dimensions may not be organized in such an orthogonal manner but in a hierarchical combination (Wickens, 1984). Each cell in the space may be conceived of as a semi-independent subsystem. Resources are shared within each subsystem but are not transferable to another subsystem of the same dimension. For instance, perceptual/central resources cannot be

exchanged to compensate for a deficit in the response resources. When two tasks compete for the same resources within a subsystem, resource competition occurs and a deficit is shown in the performance. When two tasks spread demands over separate subsystems, there is little interference and time-sharing is efficient.

-----  
Insert Figure 1 about here  
-----

Mental workload is specified as a vector in the multidimensional space presented in Figure 1. Workload is defined in terms of resource demand of a given task and competition between dual tasks on each dimension. Two types of performance-based measures can identify vectors in such a space. They are primary task performance and secondary task performance (Wickens, 1984).

Primary task performance is considered a vector measure because it directly reflects the resource demands imposed by task performance on the three dichotomous dimensions. However, there is one drawback in this measurement technique. It cannot reflect resource consumption when a process is data-limited (Norman & Bobrow, 1975). When a process becomes data-limited, performance maintains at an asymptotic level even as subjects actually invest more resources to do the task.

The resource demand of a task may also be measured by the secondary task technique. A secondary task is added to absorb the residual resources left from a primary task. Therefore, secondary task performance is inversely related to the demand of the primary task. The secondary task technique may also be used to identify a structure-specific vector, imposed by a primary task on the resource dimensions, by analyzing the interference patterns between the primary task and a set of secondary tasks. Because

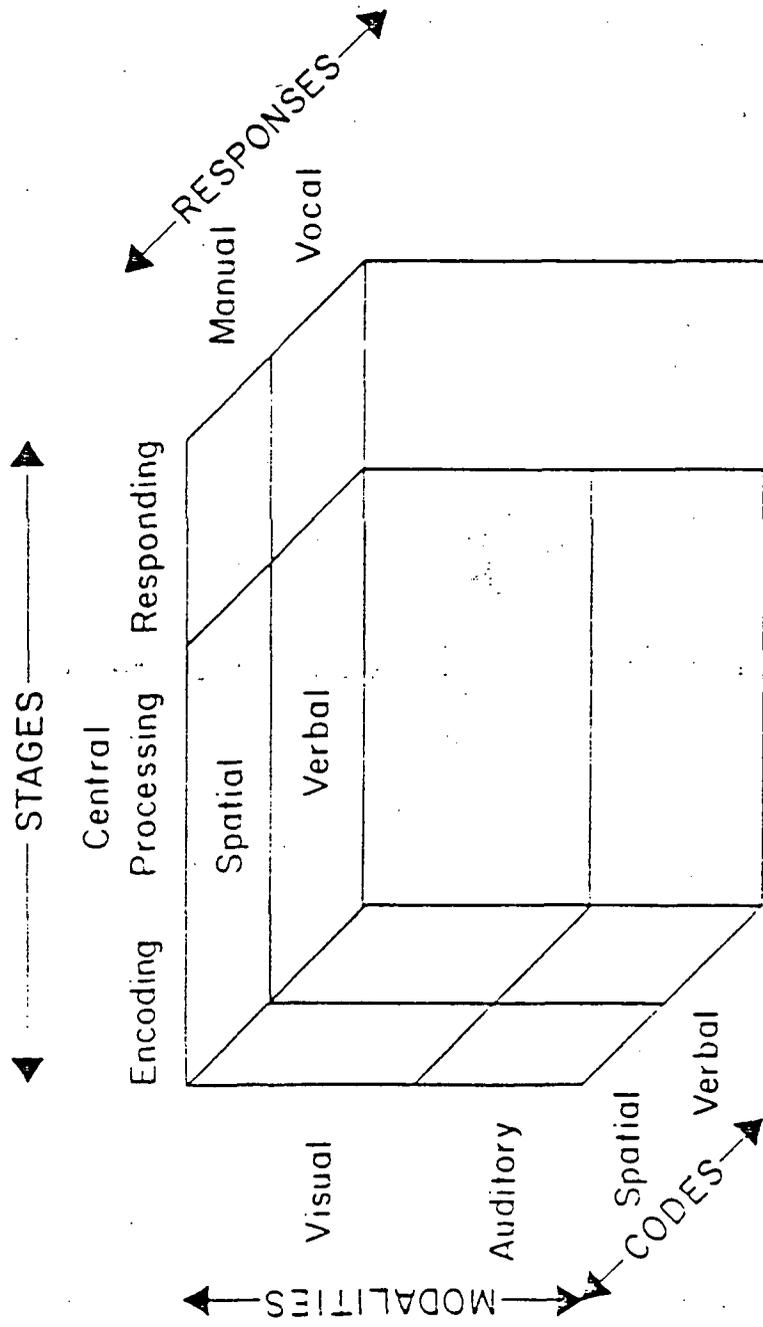


Figure 1. The proposed multiple resources model (Wickens, 1981, 1984)

secondary task performance may be used to determine the locus of resource demand of a primary task, it is considered a diagnostic vector measure.

### Subjective Workload

Although workload may be indicated by the two types of performance measures, many researchers argue that the subjective aspect is crucial to the concept of mental workload. Sheridan (1980) stated that subjective perceptions of cognitive effort constitute the essence of workload. Others suggested that "if an operator feels effortful and loaded, he is effortful and loaded" (Johannsen et al., 1979, p. 105). Many researchers take the view that if you want to measure an operator's workload, you should simply ask him what his workload is. In addition to this face validity, subjective measures are considered nonintrusive, cheap, and convenient to be implemented in any environment.

With these advantages, subjective measures are regarded as: (1) central to any investigation of workload (Johannsen et al., 1979), (2) valuable indices of workload (Moray, 1979; Wierwille & Williges, 1978), and (3) the most sensitive and reliable measures (Gartner & Murphy, 1976). Furthermore, subjective measures and performance generally show a fairly high level of correlation across a wide range of tasks and task configurations. Tasks that are performed more poorly are generally described as more difficult.

Regardless of their popularity, little effort has been made to understand the origins of subjective feelings of load (Moray, 1982) and the limitations of the techniques. Many important questions about subjective measures remain unanswered. For example, what do subjective ratings

measure? How sensitive are they to the demands on different resource dimensions? Under what conditions do they dissociate from performance?

In the present study, these aspects are examined. The structure of perceived workload is revealed and is compared with the structure of processing resources provided by the multiple resources model. Since performance and subjective workload assess the same hypothetical construct, a strong correspondence is expected. A model of verbal report data is used to understand the limitation of subjective introspection. A theory of dissociation is proposed to specify the conditions in which performance and subjective workload dissociate. Finally, results from three experiments are presented to test the hypotheses drawn from the proposed theory. It is believed that the results of the present study not only advance our understanding of subjective workload, but also indicate why performance and subjective workload dissociate.

The structure of perceived workload: A Multidimensional Scaling approach. The structure of perceived mental workload has been investigated by a Multidimensional Scaling (MDS) approach (Derrick, 1981; Derrick & Wickens, 1984; Yeh & Wickens, 1984). In this MDS approach, tasks were judged by the similarity of difficulty in order to disclose the hidden cognitive structure. Results from these studies demonstrated that the subjective aspect of workload may be understood by the dimensional structure that has been empirically verified and provided by the multiple resources model. The structure of subjective workload was found to be closely related to the structure of processing resources. The structure was composed either of three dimensions (input modality, resource competition, and time-demand

in Derrick & Wickens's study), or two dimensions (resource demand and processing codes in Wickens & Yeh's study).

Although the structure of subjective difficulty corresponds with the structure of information processing and performance, subjective measures are limited by their own nature. On the one hand, subjective ratings are scalar measures and hence do not completely preserve the characteristics of a vector. On the other hand, subjective ratings of workload are verbal report data and this imposes a potential limitation on their accuracy.

A model of verbal report data. According to Ericsson and Simon's model of verbal report data (1980), introspection reflects information heeded in working memory. Any information to be verbally reported has to be in working memory or has been transferred to long-term memory. Automatic processes that do not utilize capacity in working memory will not be available to introspection. Processes whose demands exceed the maximum capacity of the memory will not be accurately reported because there is less variation in resource mobilization under these conditions.

This model of verbal report data is consistent with previous findings in workload measurement studies. Most variables that have been found to affect subjective workload are related to the demands on working memory. These variables include memory load and presentation rate (Hauser, Childress, & Hart, 1982; Daryanian, 1980), the number of tasks-to-be-processed per processing unit (Tulga & Sheridan, 1980), the requirement of generating lead and making precise control, rate of information processing, fraction of attention allocated to the task, insufficient data, number of decision alternatives, etc. (see Moray, 1982 for a review). Moreover, Eggemeier, Crabtree, and Reid (1982) showed that subjective ratings were not

as sensitive as was performance in a high memory load condition, as suggested by Ericsson and Simon (1980).

Gopher and Braune (1983) presented a similar view and suggested that subjective measures reflect the perceived magnitude of resource investment in the conscious attention. They argued that we are only aware of part of the information processing that we do. Thus, subjective estimates follow the pattern of the most restricted model of a single undifferentiated pool of resources while performance follows the pattern of a multiple resources model.

As a consequence of this restricted source of input to subjective measures, their correlation with performance should not be expected to be unity or even fairly high across all the conditions. Human performance is determined by the interaction of the capacities of a large number different subsystems and the demands imposed on those subsystems. Wickens and Yeh (1983) pointed out that there is no reason why the demands on different subsystems must be equally read when an operator generates an introspective rating of mental workload. Dissociation arises when certain subsystems contribute very heavily to subjective workload estimates but only marginally to performance. Tasks that impose on those subsystems will be performed considerably better than their subjective ratings will indicate. Dissociation also occurs when subsystems contribute heavily to performance, but are not read by subjective measures. In this case, subjective measures will provide an overly optimistic view of the expected level of system performance.

### A theory of dissociation between performance and subjective workload

Based upon the model of subjective introspection and an examination of the sources of dissociation that have been found in the literature, Wickens and Yeh (1983) proposed a theory of dissociation from the framework of the multiple resources model. Subjective workload was proposed to be sensitive to the amount of aggregate resource investment. Furthermore, if subjective introspection reflects information heeded in working memory, primarily represented by the perceptual/cognitive resources, then the demands on these resources may contribute more to subjective measures than to performance. Several sources of increased information processing demands may produce dissociations between performance and subjective measures. These sources are listed in Table 1.

-----  
Insert Table 1 about here  
-----

Things that can be done to a task to increase the workload (decrease performance and/or increase the subjective feeling of effort) are listed as the set of manipulations in the first column. Within the second column are numbers that indicate the extent to which each manipulation will deteriorate performance (P). The numbers within the third column indicate the increase of subjective difficulty (S) produced by each manipulation. The particular values of these numbers are somewhat arbitrary, but what is important is the relative value of these two numbers, P and S, for a given source. In other words, the numbers are ordinal and only the relative order between the numbers is meaningful. The ratio or the distance of two numbers provides a qualitative description rather than an absolute quantitative value.

Table 1: A theory of dissociation

Sources	<u>P</u> (decrease)	<u>S</u> (increase)
1. Increased single task difficulty	4	3
Perceptual/Cognitive	2	2
Response	2	1
2. Concurrent task demand	3	4
Same resources	2	2
Different resources	1	2
3. Motivation	-1	+1
Resource investment		

This theory proposes that manipulating the parameters of a single task generally influences P as much or more than S. This difference is particularly pronounced if the single task is degraded by imposing demands on responding, rather than on perceptual/cognitive processing. The theory also suggests that increasing workload by increasing the number of concurrent tasks generally serves to increase S and decrease P, but the former by a greater degree than the latter. In addition, subjective experience of workload is postulated to be uninfluenced by whether those tasks compete for common or separate processing resources. But, performance suffers substantial losses when common resources are employed in both tasks and is little affected when demands are spread over separate resources. A final source of dissociation is related to any variable that induces more resource investment to improve task performance. Subjective workload is predicted to be increased by this variable even if the investment helps performance to a great extent.

This theory has been partially supported by previous studies and in fact was formulated in part to account for the data of those studies. In Derrick's study (Derrick, 1981; Wickens & Derrick, 1981; Derrick & Wickens, 1984), the difficulty of a manual control task was varied by three means: (1) increasing the control order of the task which demands both perceptual/central and response resources, (2) adding a concurrent task (an auditory memory search task or a tone judgment task) which demands non-overlapping resources, and (3) pairing the task with itself or with a visual search task which competes for common resources. All of these manipulations increased demands upon the processing system and thus induced decrements in performance and increments in subjective ratings of workload. Dissociation was found between the two measures under these manipulations. Subjective

workload was higher under an easy dual task condition than under a hard single task condition even though performance was better under the former condition.

In a study conducted by Wickens and Yeh (1983), the difficulty of a tracking task was manipulated by three means: (1) the control order or the bandwidth of a first-order tracking task was increased, (2) when performing a second-order tracking task, subjects tracked with or without a predictor, and (3) the tracking task was paired with itself or with a Sternberg memory search task whose input and output modalities were manipulated to vary the degree of resource competition between two tasks.

Replicating Derrick and Wickens' finding, subjective workload was higher but performance was better when doing two easy tasks than when doing a hard single task (Figure 2a). This result supports the theory of dissociation and the model of verbal report data suggested by Ericsson and Simon (1980). Even if there are separate resources available to perform two tasks, the perceptual/central resources are still highly demanded as an executive to coordinate the processing and the execution of both tasks. Relative to the demand of a hard single task, this "cost of concurrence" (Navon & Gopher, 1979) enhances the demand for perceptual/cognitive resources to a greater extent.

-----  
Insert Figure 2 about here  
-----

Wickens and Yeh (1983) also found a dissociation when different single task conditions were compared. (1) For an equal level of subjective workload, performance decrements were larger when performing a high-bandwidth task than when performing a second-order tracking task (Figure

ORIGINAL PAGE IS  
OF POOR QUALITY

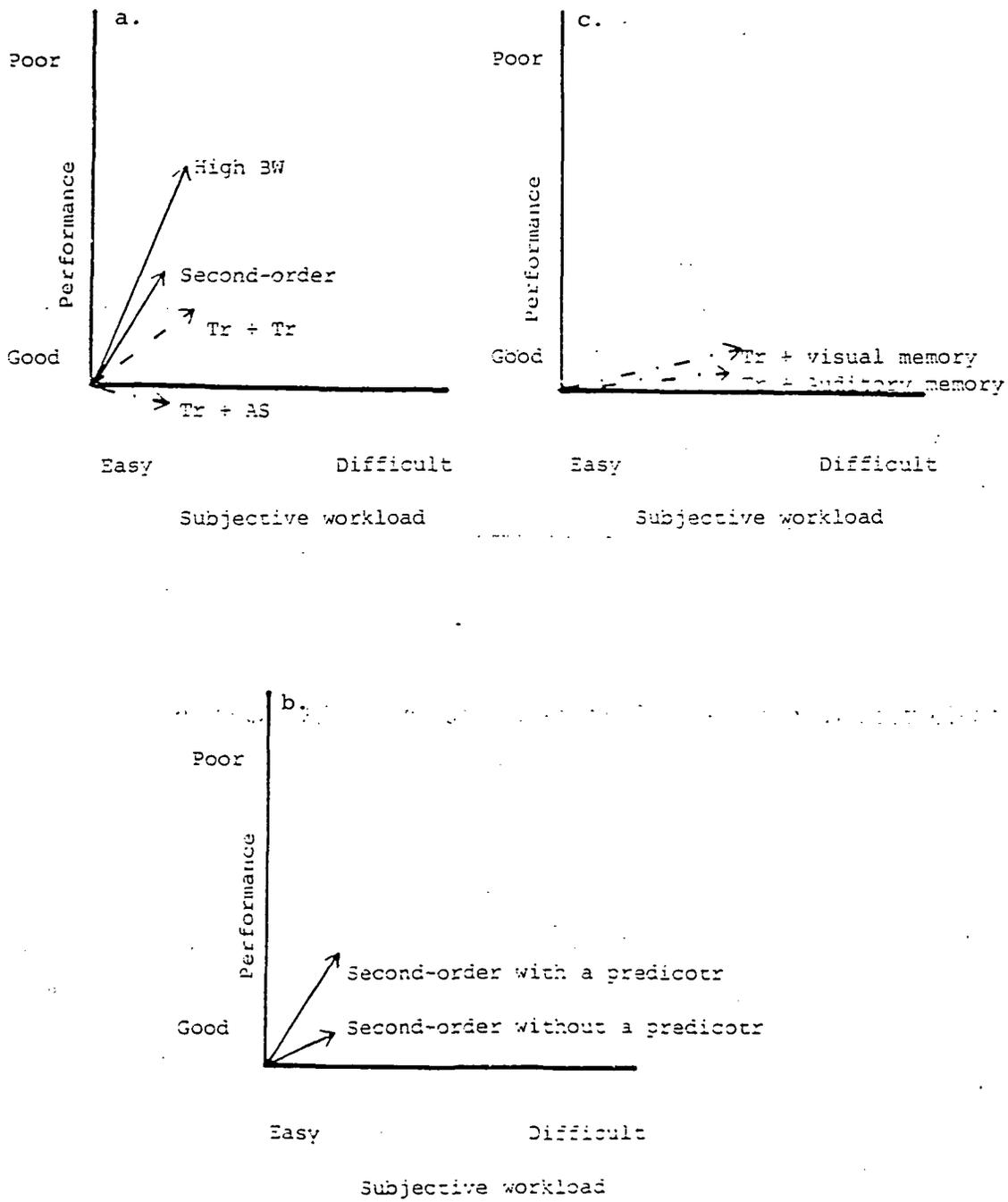


Figure 2. Dissociations found in a previous study

2a). Limited by a small sample of subjects, Wickens and Derrick (1980) tentatively suggested that a high bandwidth tracking task demanded stage-non-specific resources but a second-order tracking task demanded both perceptual/cognitive and response resources. Therefore, a second-order tracking task may demand more perceptual/cognitive resources than does a high-bandwidth tracking task. As a consequence of these different demands on perceptual/central resources, subjective workload overestimated the performance decrements in a second-order tracking task. (2) A predictor display in a second-order tracking task aided performance by providing a precise control information. Nevertheless, in order to utilize the information, subjects had to invest more resources to process two task elements (i.e., cursor and predictor). Thus, subjective workload under a predictor display condition was higher than what its performance would indicate (Figure 2b).

In Wickens and Yeh's study (1983), dissociation was also found when different dual task conditions were compared. For an equal performance level, subjective workload was higher in a low resource competition condition (tracking and an auditory-speech memory task) than in a high competition condition (two tracking tasks) (Figure 2a). The insensitivity of subjective measures to resource competition, once again indicated that some information was not read equally by the two measures. The executive management was equally engaged in coordinating two tasks and the aggregate resource investment was similar in both high and low resource competition conditions. Since the aggregate investment contributed more to subjective measures than to performance, subjective workload was not as sensitive to the degree of resource competition as was performance.

## Experimental Overview

Although the theory of dissociation proposed by Wickens and Yeh (1983) has been partially supported, a systematic approach was adopted in the present study. Converging evidence can provide a thorough test of the theory and can advance our understanding of subjective workload. Sources of information processing that may produce dissociation were directly manipulated and hypotheses drawn from the theory were tested. Three tasks: a Sternberg memory search task, a one-dimensional compensatory tracking, and a dual-axis tracking task were used in three experiments. In the Sternberg memory search task, demands were imposed upon different stages of processing or different codes of central processing. In the one-dimensional compensatory tracking task, the bandwidth or the control order was varied to create the difficult conditions. Display augmentation was implemented to aid performance in two higher order tracking task conditions. In the dual-axis tracking task, the effect of separating display or control was investigated. In dual task conditions, the degree of resource competition was manipulated via various combinations of the tracking tasks and different Sternberg memory search tasks. The patterns of dual task interference were analyzed by the additive factors methodology in order to make inferences concerning the resource demands of various tracking tasks. The following section provides a rationale for the specific manipulations that were chosen for testing each hypothesis.

## Hypotheses of the dissociation

Hypothesis 1 : Subjective ratings are dominated by the demands on perceptual/central resources. According to the theory, dissociation will occur when two manipulations increase their demands on different processing stages. In the first experiment, demands were placed on a specific processing stage (i.e., perceptual load, central load, or response load) to increase the difficulty of a Sternberg memory search task. Each difficulty manipulation was assumed to effect the amount of demands on one stage without altering the demands on other stages. If the hypothesis is correct, then dissociation should be found when the perceptual or central load memory task was compared with the response load condition. For an equal performance level, subjective workload should be greater under the perceptual and central load conditions than under the response load condition.

The integrality of a cursor in a dual-axis tracking task was manipulated in the third experiment. In one condition, two cursors were tracked by one joystick to impose extra demands upon perceptual/central resources. In another condition, one cursor was controlled by two joysticks to increase the demands on response resources. A dissociation was predicted when these two dual-axis tracking tasks were compared.

Hypothesis 2: Subjective ratings are driven more by the number of concurrent tasks whereas performance is relatively more effected by the difficulty of single tasks. This hypothesis was postulated to replicate previous findings in different task configurations. In the first experiment, the difficulty of a single Sternberg memory search task was

increased by enhancing the demands on specific processing stages. The difficulty of a tracking task was manipulated by increasing the bandwidth. The difficulty was also increased by combining the two easy tasks in a dual task condition. In the second experiment, the difficulty of a tracking task was raised by increasing the bandwidth, the control order, or by adding a basic memory search task in a dual task condition. In the third experiment, the difficulty of a dual-axis tracking task was varied through the demands on perceptual or response resources in single task conditions. The difficulty was also increased by combining the tracking task with an easy memory search task in a dual task condition.

The number of display elements affects subjective workload as much or relatively more than it influences performance. A related test of the first two hypotheses is that the number of display elements drives subjective workload more than it effects performance. Murphy, McGee, Palmer, Paulk, and Wempe (1978) showed that when information presented in the center of a flight control display was cluttered with a number of motion indicators, performance was better but subjective workload was higher in this condition than in the other two display conditions in which fewer motion elements were visible. Garner (1974) showed that integral and separable dimensions influence subjective perception. Perceptual features are perceived as one unit when they interact integrally and as distinct units when they are separable. Kahneman and Triesman (1984) proposed a similar concept (i.e., the notion of "object file"). Features of an object are perceived as a whole and attention is directed by the object rather than by the composite features. When there is more than one object, attention has to be divided

between objects. Kramer (1984) also showed that the number of object units effected dual task performance.

Wickens and Yeh (1983) found that an external predictor in a second-order tracking display improved performance but did not decrease subjective workload. In interpreting the result, they suggested that the external predictor was separate from the tracking cursor and was hence perceived as another unit in working memory. Resource demands on working memory under the predictor display condition were higher than the demands of a tracking task in which the external predictor was absent. Therefore, subjective workload was higher under the former condition than what its performance would indicate.

The integrality of a predictor with a tracking cursor was manipulated in the second experiment. In one second-order tracking condition, the predictor was external to the cursor. It was assumed that processing two distinct elements, in comparison with a display without the predictor, would demand extra perceptual/central resources, but that the predictor symbol would nevertheless improve performance. In another condition, the predictive information was integrated with the cursor so that only one object unit was processed. If the number of object units affects the perceived demands in addition to the objective demand, a dissociation will occur. For an equal performance level, integrated features will be perceived as less demanding than separate features.

Hypothesis 3 : Subjective ratings are not as sensitive as is

performance to resource competition. If subjective ratings measure the aggregate demands, ratings should be roughly equivalent in both high and low resource competition conditions but performance will depend upon the amount

of resource competition. In previous studies (Derrick & Wickens, 1984; Wickens & Yeh, 1983), this dissociation was shown when the difference in the amount of competition was large between two dual task conditions (dual tracking tasks vs. tracking with an auditory memory task).

In the present study, the amount of resource competition between two tasks was restricted to certain resources. In Experiment 1, the degree of competition for specific stage-related resources was varied in different dual task conditions. A tracking task presumed to be "response loading", was performed concurrently with a response load memory task (high competition) or with a perceptual/central load memory task (low competition). The competition for the codes of central processing was manipulated in the second and third experiment. A tracking task was combined with a spatial memory task (high resource competition) or with a verbal memory search task (low competition). It was predicted that dissociation would occur when the low and high resource competition conditions were compared. Performance decrements would be larger in the high competition conditions but subjective workload would be relatively less sensitive to the difference in the amount of resource competition.

Hypothesis 4 : Factors which induce more resource investment to improve performance will also increase subjective workload. Moray (1982) suggested that the degree of precision required by a task and the probability of failure affect subjective mental load. The degree of precision in a high-bandwidth tracking task was used in the second experiment to motivate subjects to invest more resources to improve performance. It was predicted by the theory that this investment would also increase subjective ratings of workload and dissociation would occur.

Tulga (1978) argued that subjects may lower the performance criterion in a high load condition when they perceive a large discrepancy between the desired and the actual performance. When subjects lower their criterion, their performance will drop off but they will feel less loaded. To test this potential source of dissociation, two degrees of precision were imposed upon subjects in different high-bandwidth tracking task conditions in the second experiment. A very precise control was required in one condition so that the subject's actual performance would be far from the objective demand. In another condition, the precision was eased so that the perceived discrepancy between the actual and the objective performance would be relatively smaller. If subjects lower their performance criterion in the former condition, dissociation between the two workload measures will occur.

#### Secondary issues

In addition to verifying the theory of dissociation, several secondary issues were also investigated in the present study. These issues include:

The resource demands of a high-bandwidth tracking task and of a dual-axis tracking task with separated display/control. Wickens and Derrick (1980) tentatively suggested that the bandwidth manipulation requires stage-non-specific resources. Manipulations similar to those used in the preceding experiment were employed to verify this suggestion. In the first experiment, the bandwidth manipulation was combined with the manipulation of Sternberg variables on specific processing stages in various dual task conditions. The locus of resource demand of the bandwidth manipulation was analyzed by the additive factors method. The bandwidth manipulation was combined with the manipulation of processing codes in the second experiment

to test whether increasing bandwidth places demands on a specific code of central processing. The same method was also applied to investigate the locus of resource demand of a dual-axis tracking task with display or control separated.

The reliability of subjective ratings. Although subjective measures have been widely used, the reliability of these ratings has seldom been investigated. Yeh and Wickens (1984) showed that post trial ratings were less reliable than post session ratings even though both were fairly reliable. In the present study, this issue was examined to replicate such a difference in rating reliability because the implication of this issue is important. If operators are more consistent in rating when they have a chance to compare the demands of all the systems, ratings should be collected in the context of all of the just-performed systems. Ratings collected immediately after operators test a system may be less reliable.

Predicting a global workload scale from specific scales. Using a multiple regression analysis, the rating on a global workload scale may be treated as the dependent variable and the ratings on other scales may be treated as the predictors. The multiple R square indicates the amount of variance in the global workload rating that is accounted for by the predictors. The regression weight of each predictor represents its importance in predicting the global rating. If the weight is not significant, then this predictor is not a salient component of the global workload measure. Vidulich and Wickens (1984) found that only three specific scales contributed to explain the variance in the ratings of two global scales. These scales are mental/sensory effort, response load, and

stress level. Two of the scales are directly related to the resource dimensions described by the multiple resources model.

The perceived structure of subjective workload. Results from previous studies (Derrick & Wickens, 1984; Yeh & Wickens, 1984) demonstrated that there is a strong relation between the perceived structure of subjective difficulty and the structure of processing resources. The multidimensional scaling approach was again adopted in the present study to disclose the hidden cognitive structure which may not be in the conscious awareness of the subjects. This evidence may provide a complete picture of the structure of subjective workload. Furthermore, the nature of subjective workload revealed from the MDS analysis may be combined with the results of the regression analysis to help the understanding of the performance-subjective workload dissociation.

#### Methodology of the analysis

An analysis method is necessary to test whether different manipulations have differential effects on the performance and subjective ratings. The interaction effect of an ANOVA provides such an information. Prior to the analysis, performance and subjective workload measures must be scaled into common units in order to test the relative effects of two manipulations. However, performance and subjective measures differ in their nature. Subjective ratings are scalar measures whereas performance is itself a vector quantity, sometimes reflecting one, and sometimes two tasks, and having speed and accuracy components. Performance from each facet must be integrated into a scalar measure to reflect the total effect of a

manipulation. In addition to providing the dependent variable for the test of dissociation, this integrated performance measure may be used to examine the interference pattern between two tasks. The interference pattern is used to infer the processing demands of various tracking tasks via the additive factors methodology. Each of these points is discussed in the following section.

#### Test of dissociation

Assuming that the two measures are expressed in comparable units, a dissociation may be characterized in a space with subjective ratings of workload and performance measures as the two axes. Ideally, if a manipulation drives the two measures to the same degree, its effect as measured from a baseline condition will result in a 45 degree vector in the space. Both the ratio of change in performance to change in subjective measures ( $\Delta P/\Delta S$ ) and the correlation between the two measures will be 1.0 (Vector a in Figure 3a). When a manipulation effects one measure relatively more than the other, the vector will deviate toward one of the axes (Vector b in Figure 3a).

-----  
Insert Figure 3 about here  
-----

Increases in demands from a baseline condition (e.g., a first-order low bandwidth tracking task or a Sternberg task with memory set of one) may be imposed by different means. When different manipulations drive the two measures to the same degree, the resulting two vectors will have the same slope (Vectors b and d in Figure 3a). When manipulations drive the two measures to different degrees, the resulting vectors will diverge (Vectors b and c in Figure 3b). A significant divergence between two vectors indicates

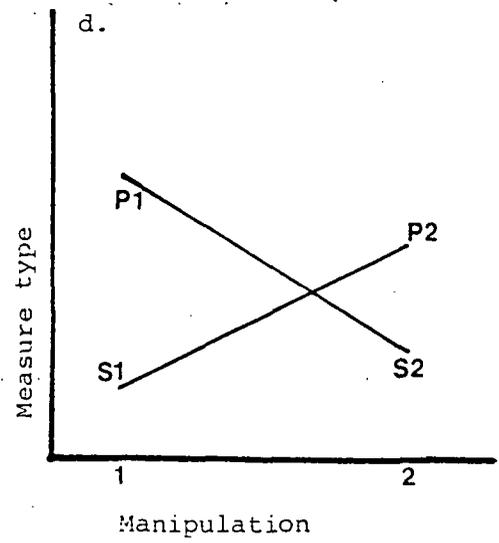
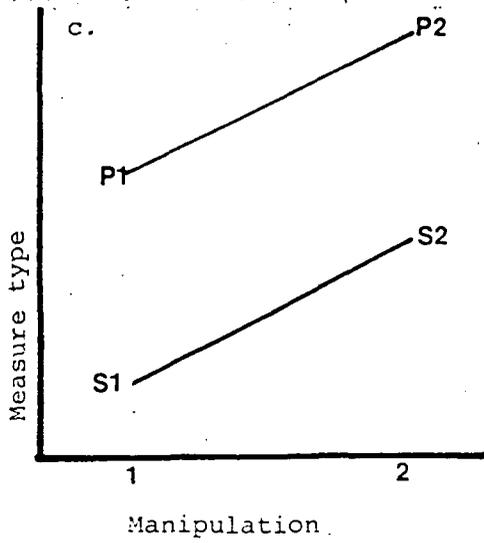
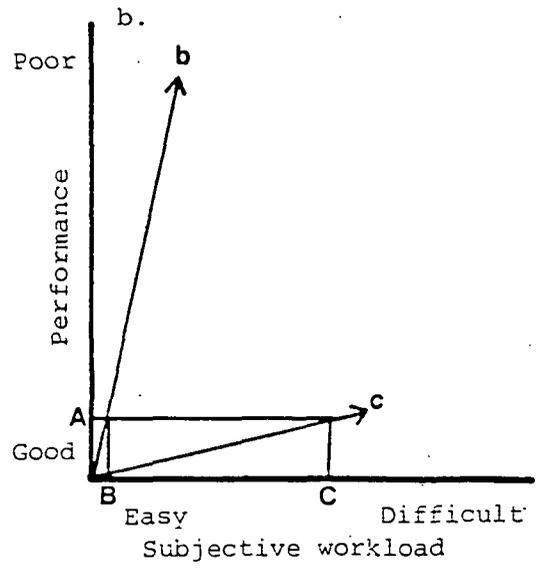
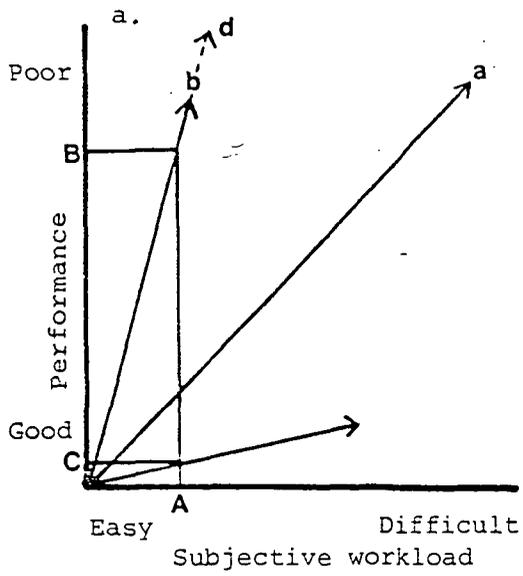


Figure 3. A hypothetical dissociation and the test of dissociation

a dissociation between the two measures as the consequence of different manipulations.

With the baseline condition(s) represented as the origin of the vectors, a dissociation is defined by two characteristics. (1) For an equal level of subjective workload (Point A in Figure 3a), there is a significant difference between the performance measures (Point B vs. C in Figure 3a). (2) For an equal performance level, there is a significant difference between subjective ratings of workload (Figure 3b). If there is no significant dissociation between the two measures, then P1/S1 will be the same as P2/S2 and there will be no interaction between the measure type and the manipulation type (Figure 3c). If there is a significant dissociation, then the measure type will significantly interact with the manipulation type (Figure 3d). Therefore, a significant interaction in the ANOVA, with the manipulation type and the measure type as two factors, is used to indicate diverging vectors in the performance-subjective ratings space.

#### Workload measures in comparable units

How should different types of measures be scaled in comparable units in order to test a dissociation? Subjective measures are usually assessed by certain rating values (i.e., interval or ratio scales). On the other hand, performance measures are recorded in various forms such as latency, accuracy, or tracking error. There are three transformation methods which may be employed to produce comparable units.

(A) Standardized scores. In this method, a grand mean and a standard deviation are obtained from all the subjects and all the conditions for each dependent measure. Each subject's score is then standardized by the grand mean and the standard deviation. This procedure is executed for all facets

of performance and for subjective measures. Various performance variables are then weighted and combined to produce one performance measure in each dual task condition. After this transformation, both performance and subjective measures are on the standard units.

(B) Estimates of effect magnitude. For each dependent measure, an effect size is derived for each difficult condition (see Hedges, 1981, 1982; Ackerman, in press for a complete discussion). Basically, the effect size is the t-value obtained by (1) subtracting the group mean of the control condition (easy single task) from the group mean of the experimental condition (e.g., a dual task condition or a hard single task), and (2) dividing this difference by the weighted variance of the two conditions. The effect sizes (t-values) of the dependent measures are then combined to estimate the overall magnitude of effect in a dual task condition or a hard single task condition. The subjective measures and performance are now on a common metric of effect magnitude.

(C) Decrement scores. To obtain the effect of one manipulation on a control task, a difference score is first derived for each dependent variable and for each subject. For each measure, the variance in the difference scores is computed across all the subjects for each condition and the mean variance across all the conditions is used as the denominator to normalize the difference scores. This method is a modified version of the one used by Wickens, Mountford, and Schreiner (1981). The normalized decrement score combined between two tasks is then used as the performance measure of a dual task condition.

In essence, both of the latter methods attempt to measure the size of an effect relative to the control conditions. However, the decrement measure (Method c) differs from the method of estimating effect size in the

derivation of variances. In Method c, individual decrement scores rather than the raw scores are used to obtain the variance of each condition. These individual decrement scores are used so that the effect size of each subject's performance or the effect size of the rating is not influenced by other subjects' effect sizes. These measures are particularly appropriate for subjective ratings because different subjects use the rating scales in various ways. Subjects vary widely in the values and the ranges given to tasks. In addition, the variance is averaged across all the conditions in Method c, instead of the two conditions that are of interest in a comparison. The basic assumption underlying this averaging is that the decrements in all of the conditions are drawn from one distribution.

#### Test of the resource demand of a task: The additive factors methodology

The additive factors method was originally employed to analyze the latency of a mental process into distinct stage components (Sternberg, 1966). One manipulation factor is assumed to prolong the duration of a particular stage without altering the latency of any other stage. Since stage durations are additive by definition, the changes in mean RT produced by two factors will be independent and additive if they effect different stages. When two factors influence the same stage, then effects of the two on RT will be interactive. This method has been applied in dual task researches to localize the processing resources that overlap between a manipulated Sternberg task and a primary task (Wickens, 1978; Micalizzi & Wickens, 1980). When the Sternberg manipulation and the primary task manipulation consume the same resources, the increase in latency of both manipulations together will be greater than the sum of the independent

increases. When the two demand non-overlapping resources, additivity occurs.

In the present study, the dependent variables characterizing the performance of a single memory task were the average of the latency and the accuracy under that condition. The dependent variable of a dual task condition was a score combined from the performance of each component task (i.e., RMS, RT, and accuracy). It was assumed that the effect of a manipulation on resource dimensions may be manifest in all the facets of performance. Therefore, the single latency dependent variable may not truly reflect the effect of a manipulation because subjects may adopt a tradeoff strategy. For example, subjects may trade accuracy for a faster reaction time in single memory task conditions, but not in dual task conditions. Or, they may protect memory task performance in a dual task condition by trading the tracking performance. Thus, it was assumed that the combined performance measure reflects the total cost imposed by the concurrent resource competition, independent of resource allocation between two tasks.

Therefore in the present study, after being transformed to comparable units, performance measures from all facets of processing are integrated. The integrated performance measure represents the total effect of a manipulation on the human information processing system. This integrated measure, employed with the additive factor methodology, is used to test for the interference patterns of various dual task conditions. The results are used to infer the demand of (a) a high-bandwidth tracking task, (b) a second-order tracking task with a predictor, and (c) a dual-axis tracking with display or control separated. These various tracking tasks were combined with different Sternberg-task manipulations in dual task conditions across the three experiments. When a Sternberg-task variable and a tracking

difficulty manipulation demand the same resources, the effects of the two on the integrated performance measure are assumed to be interactive. When the demands of the two manipulations are spread over separate resources, the effects are assumed to be additive.

In the following section, the results of three experiments are presented in concert because the data are related to the same issues and test common hypotheses.

## Experiment 1

Various combinations of a tracking task and a Sternberg memory search task were employed to investigate three hypotheses. To review, these hypotheses state that (1) Demands upon the perceptual/central resources drive subjective ratings more than they drive performance. On the other hand, response complexity drives performance more than it drives subjective ratings of workload. (2) Performance measures rather than subjective ratings are sensitive to resource competition. For an equal level of performance, subjects feel more loaded when performing a low competition dual-task combination than when performing a high competition condition. (3) Dissociation occurs when a factor induces more resource investment to improve performance.

The difficulty of a single memory search task was increased by imposing more demands upon a specific processing stage. The manipulations resulted into three hard single task conditions (i.e., perceptual, central, and response load memory tasks). Performance and subjective measures of these three conditions were then contrasted to test the first hypothesis. The bandwidth of a tracking task was increased in the difficult condition. Manipulations of Sternberg variables and of tracking bandwidth were combined in dual task conditions to (1) infer the resource demand of a high-bandwidth tracking task, and to (2) test the second hypothesis on the effect of resource competition. To test the third hypothesis, the degree of precision in a high-bandwidth tracking task was manipulated as a factor that motivates subjects to invest more resources.

## Method

### Subjects

Fifteen students of the University of Illinois participated in this experiment. All subjects were right-handed, native speakers of English, and had normal vision.

### Apparatus

Subjects were seated in a light and sound attenuated chamber. The tasks were implemented on a PDP-11/40 computer. The computer was interfaced to a 10 cm x 8 cm CRT display via a Hewlett-Packard 1300 Graphics Display Interface. The display was about 90 cm in front of the subject and slightly below eye level. The subject's responses for a memory search task were accomplished through a push button control panel affixed to the right armrest of the subject's chair. The subject's input for a manual control tracking task was via a MSI 521 joystick affixed to the left armrest of the subject's chair.

### Tasks

Sternberg memory search task. Prior to each trial of the basic memory search task (BM), the subject viewed one alpha-numeric string (e.g., A26). The string was held in working memory for the next two minutes as the subject was presented a series of probes, half of which were in the memory set. Subjects were asked to press the first key on the panel if a probe was in the memory set (i.e., a "go/no-go" response). The demand on each processing stage was increased via different manipulations (Derrick & Wickens, 1980). In the perceptual load condition (PM), a grid mask was

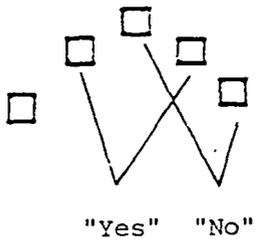
superimposed upon each probe stimulus. In the central load condition (CM), memory set size was increased to three (e.g., A26, C59, R89). In the response load condition (RM), subjects had to make double responses for every positive response and for every negative response. Subjects had to press key 2 then key 4 for a positive response. They had to press key 3 then key 5 for a negative response (Figure 4a). Subjects had to release the first key completely before they pressed the second key and they had to press the second key within a window of 600 msec. Responses outside this window were recorded as non-responses. Reaction time of the first response was used as the latency of the response load condition. The subject's task in all of these conditions was to respond as accurately and as quickly as possible.

-----  
Insert Figure 4 about here  
-----

Tracking task. In the easy condition, subjects were required to nullify an error cursor that was displaced horizontally by a random forcing function with a cutoff frequency of 0.32 Hz. The error cursor was controlled by the left-right movement of the joystick with the left hand. Only the first order system dynamics were utilized. The bandwidth was increased to 0.54 Hz in the difficult condition.

When subjects tracked the cursor in a "high-bandwidth with a window" condition (Figure 4b), an acceptable error range was presented. This manipulation was employed to motivate subjects to invest more resources to improve performance. Subjects were told that this window indicated a safe altitude level and that they should keep the cursor within the window as much as possible in order to keep the control of the plane. An auditory

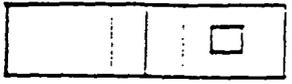
A29



High BW tracking



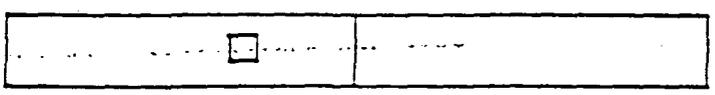
Wide window



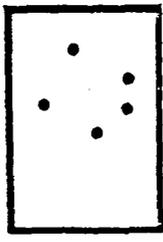
Narrow window

a. Response pattern in the response load condition

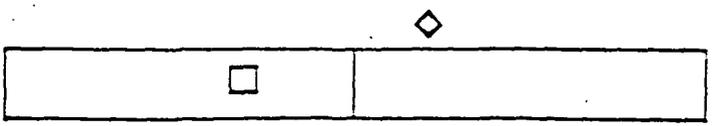
b. High-bandwidth tracking



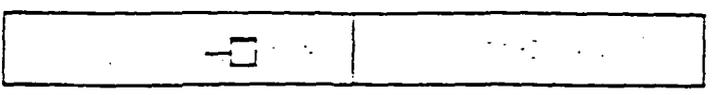
No augmentation



c. An example of a spatial memory set



Predictor display



Command display

d. Second-order tracking

Figure 4. Response pattern in the response load condition and some displays seen by the subjects

tone was also given when the cursor was outside the range until subjects tracked the cursor back into the window. The width of the window was also varied (i.e., narrow and wide window conditions) to investigate the effect of the perceived discrepancy between the desired and the actual performance.

Dual task. Dual tasks were generated by combining one of the four Sternberg memory search tasks (i.e., basic, perceptual load, central load, and response load) with a low-bandwidth or the "no-window" high-bandwidth tracking task. Subjects were instructed to give equal priority to each task.

### Procedure

Practice sessions. Subjects practiced all of the tasks extensively in the first two sessions over a period of two days. A histogram of RMS error data was obtained from each low-bandwidth tracking task condition. These data were averaged for each subject to represent his average performance under the easy tracking condition and were used to compute the window width employed in the experimental sessions. The width was chosen so that 15% of an individual subject's tracking error distribution would fall in the "narrow" window and 30% would fall in the "wide" window range.

Experimental session, Session 3. Prior to this session, subjects were given a list of all the rating scales (the descriptions are listed in Table 2). The experimenter went through the list with the subjects and made sure that they understood the meaning of the scale and the way to do the rating. Subjects performed the basic memory search task first and they were instructed to use the subjective experience of doing this task as the standard which was arbitrarily assigned a rating of 100. Subjects then performed the other tasks and rated each one against the standard on eight

rating scales. They were told to assign a rating of 200 to a just-performed task if the subjective experience was twice the experience of doing the standard task and to assign 50 if it was half. This strategy was not adopted for the "excess capacity" scale on which subjects were allowed to assign any number between 0 and 100 for every task.

-----  
Insert Table 2 about here  
-----

The unidimensional ratings were collected at the end of each trial immediately after subjects performed the task and at the end of the session. Subjects were then given 20 pairs, randomly chosen from possible paired combinations of 14 tasks (the two window conditions were not included), and were asked to rate each pair on a 9-point similarity scale (Table 2).

Experimental session, Session 4. Subjects followed the same procedure as in Session 3 except that all the 91 possible paired combinations of the 14 tasks were rated on the 9-point similarity scale. Unidimensional ratings collected in this session were compared with the corresponding ratings recorded in Session 3 to estimate the reliability of the ratings. The similarity judgments of the 20 pairs collected in Session 3 were collated with the corresponding ratings in this session to compute the reliability of the similarity ratings.

## Experiment 2

The purpose of this experiment was to verify two hypotheses: (1) subjective workload is affected by perceptual features, and (2) subjective

Table 2 - Description of the scales used in Experiment 1

Scale	Description
Overall workload	(WK) - The total experience during the task, including how much attention that you paid and how difficult the task was.
Perceptual effort	(PE) - How much effort did you make in perceiving the stimuli presented to you in the last task? (e.g., seeing, looking, scanning etc.)
Mental effort	(ME) - How much effort did you make in the mental activity (e.g, monitoring, thinking, deciding, planning, remembering, understanding etc.)
Response effort	(RE) - How much effort did you make in responding during the last task?
Time pressure	(TI) - How much time pressure did you feel in doing the last task? How busy were you?

Table 2 (continued) - Description of the scales used in Experiment 1

Scale	Description
Stress	(ST) - How anxious/worried/uptight/harassed or calm/tranquil/placid did you feel during the last task?
Performance	(PR) - How successful was your performance?
Overall effort	(OE) - What was the overall effort that you put in the last task?
Excess capacity	(EX) - Assuming that all the effort capacity you have is 100. How much was left in doing the last task. You may use any number between 0 and 100.

The rating scale for judging the similarity of task difficulty

1	2	3	4	5	6	7	8	9
Very Similar		Similar		Inter-mediate		Dissimilar		Very Dissimilar

measures are less sensitive to the competition between the codes of central processing.

In order to test the first hypothesis, the integrality of a predictor with the cursor was manipulated. It was assumed that integral features would be perceived as a unit whereas separable features would be processed as distinct units. If the number of object units affected subjective workload, subjects would feel less loaded when perceiving integral features than when processing separable features even if performance measures were roughly equivalent.

To test the second hypothesis, the degree of resource competition was manipulated in dual task conditions. It was presumed that resource competition would be higher when a tracking task was performed with a spatial Sternberg memory search task than when it was time-shared with a verbal Sternberg memory search task. Performance was predicted to be sensitive to the difference in the amount of competition, but subjective workload would be affected by the aggregate demand.

## Method

### Subjects

Fifteen students of the University of Illinois were subjects in this experiment. All subjects were right-handed, native speakers of English, and had normal vision.

### Tasks

Sternberg memory search tasks. In the verbal memory task (BM), subjects remembered and responded to three sets of alpha-numeric strings as

in the central load condition in Experiment 1. In the perceptual memory task (PM), subjects held two sets of strings in memory and a grid mask was placed upon every probe. In the spatial memory task (SM), subjects remembered and responded to two sets of dot patterns. The dot patterns were chosen from the ones used in Sandry and Wickens' study (1982) (see Figure 4c for an example). The go/no-go response assigned to the memory tasks in the first experiment was employed in all of these tasks.

Tracking task. In addition to the low-bandwidth and the "no-window" high-bandwidth tracking tasks employed in the first experiment, three additional kinds of tracking tasks were added in this experiment. Figure 4d illustrates the displays seen by the subjects in these three conditions. In the unaided second-order tracking task, acceleration dynamics were employed. In the "predictor display" condition, a predictor symbol, driven by the estimate of the cursor's current velocity and acceleration, was displayed above the tracking rectangle. In the "command display" condition, the principle of the "pseudo-quicken" display (see Gill, Wickens, Donchin, & Reid, 1982 for the specifications of the method) was adopted. A line segment was added to the cursor to indicate the direction of the required control input by providing information concerning the higher derivatives of the system state. The direction of the segment changed when reversal should be implemented in control direction as specified by the theoretical switching line. Intuitively, the direction of the segment was the same as the subject's control movement when an undershoot needed to be corrected. The direction switched when the system velocity, relative to the system position, would produce an overshoot if reverse movements were not made immediately. In all of the three second-order tracking tasks, the gain of the system was lower than that in the other tracking conditions.

Dual task. The low-bandwidth or high-bandwidth tracking task was time-shared with the verbal memory search task or the spatial memory search task. The perceptual memory task was performed concurrently with each of the three second-order tracking tasks.

### Procedure

Subjects practiced the tasks in the first two sessions. Data were collected in Sessions 3 and 4. The low-bandwidth tracking task was used as the standard task to rate other just-performed tasks. Most of the rating scales were the same as those used in the first experiment with the following changes: (1) the perceptual effort and mental effort scales were combined into one scale, (2) the performance scale and the overall effort scale were not used in this experiment, and (3) a task complexity scale and a modified Cooper/Harper scale were added (see Table 3 for the description of these new scales). The task complexity scale was chosen from one of the categories proposed by Sheridan and Simpson (1979). The modified Cooper/Harper scale was added because the original scale has been used in many studies and Wierwille and Casali (1983) have validated and recommended this modified version for assessing overall workload. The ratings collected in the two experimental sessions were used to compute the rating reliability.

-----  
Insert Table 3 about here  
-----

At the end of Session 3, fifteen pairs drawn from the combinations of 11 tasks (the four conditions that involved the perceptual load memory task were not included) were rated on the similarity scale. At the end of

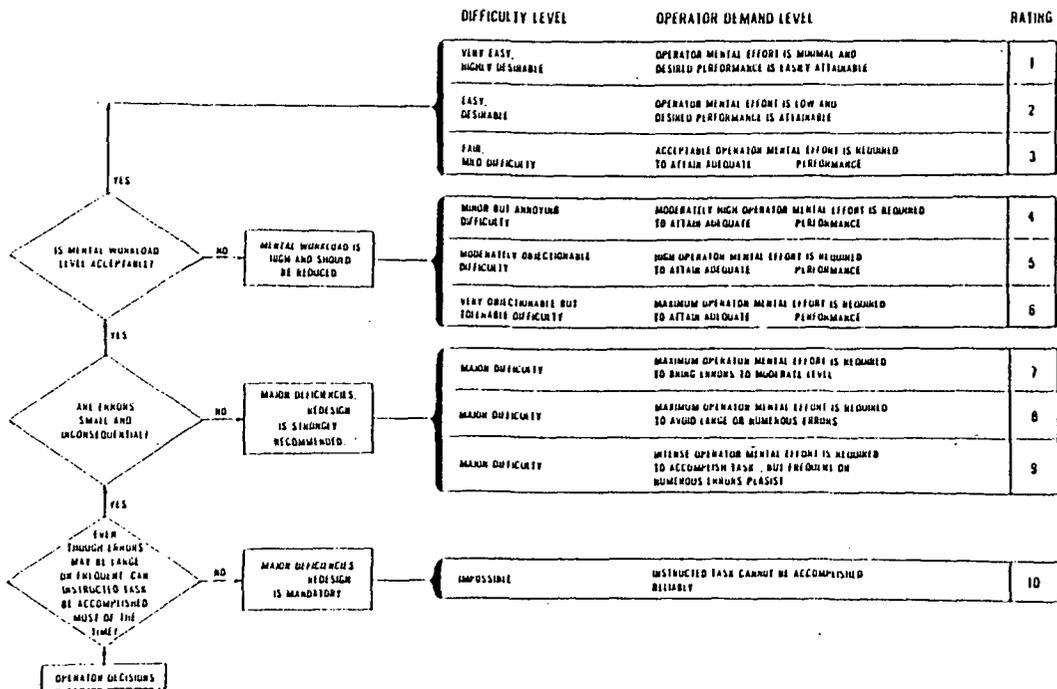
Table 3 - Description of the scales used in Experiments 2 and 3.

Scale	Description
Overall workload	(WK) - same as the scale used in Experiment 1
Perceptual/Mental effort	(P/ME) - How much effort did you make in perceiving the stimuli presented to you and in the mental activity (e.g., seeing, looking, scanning, monitoring, listening, thinking, deciding, planning, understanding, etc.)
Response effort	(RE) - same as the scale used in Experiment 1
Time pressure	(TI) - same as the scale used in Experiment 1
Stress	(ST) - same as the scale used in Experiment 1
Task complexity	(CM) - How complex was the task due to uncertainty, unpredictability, unfamiliarity, or due to automated versus skilled planning required?
Capacity demand	(CA) - Assuming that all the effort capacity you have is 100. How much did you use in doing the task? You may use any number between 0 and 100.

Table 3 (continued) - Description of the scales used in Experiments 2 and 3.

Scale Description

Cooper/Harper - Please rate the task by going through the following decision tree. It is important that you go through it step by step. You should start from the operator decision. Based on your answer to that question, you either go up to ask yourself the next question or go to the leaf for "no" answer. Do this until you choose your rating.



Session 4, all of the 55 possible paired combinations of the 11 tasks were rated on the similarity scale. The similarity ratings of the same 15 pairs in the two sessions were then collated to check the reliability of the similarity ratings.

### Experiment 3

The purpose of this experiment was to generalize the findings from the first two experiments by employing various combinations of a Sternberg memory search task and a dual-axis tracking task. In the basic dual-axis tracking task (BT), errors on the X and Y axes were integrated into one cursor on the display and were controlled by one joystick. The difficulty of this task was manipulated by imposing demands on stage-specific processing resources. It was assumed that separating the integrated cursor into two distinct indicators, with each cursor representing the error on one axis, would increase demands for perceptual/cognitive resources. On the other hand, controlling an integrated cursor by two separate joysticks would presumably demand more response resources. The difficulty of the tracking task was also increased by adding a Sternberg memory search task which demanded common or separate resources from the tracking task. The pattern of interference between the Sternberg and the tracking tasks could be used to check the locus of processing load of the two manipulations.

## Method

### Subjects

Fourteen students of the University of Illinois were subjects in this experiment. All subjects were right-handed, native speakers of English, and had normal vision.

### Tasks

Dual-axis tracking task. The objective of the basic task (BT) was to keep a cursor changing in two dimensional coordinates on the stationary target, a cross in the center, by controlling a single joystick. The dynamics of the tracking control were first order and the upper cutoff frequency was 0.30. In the "separated-display" condition (DT), two cursors signified the errors on each axis respectively and were tracked by one joystick. In the "separated-control" condition, subjects tracked one integrated cursor by moving two separate joysticks, one moving the fore-aft direction and the other in the left-right direction.

Memory search task. In the basic verbal memory task (BM), subjects remembered and responded to two sets of alpha-numeric strings. In the spatial memory task (SM), they held two sets of dot patterns in working memory and made a positive response when a probe was in the memory set. In the central load condition (CM), subjects had to remember and respond to four sets of alpha-numeric strings. In all of these conditions, a manual go/no-go response was required to minimize the response selection load. In the speech condition (SP), they remembered two sets of strings and said "yes" when a probe was in the memory set.

Dual tasks. Each of the four memory tasks was combined with each of the three dual-axis tracking tasks.

### Procedure

Subjects practiced all the tasks in the first two sessions. The basic dual-axes tracking task was used as the standard to rate other tasks in Sessions 3 and 4. The rating scales were the same as those used in Experiment 2. Similarity judgment data were also collected in Sessions 3 and 4 to compute the rating reliability.

## Results

### The processing demands of various tracking tasks

Experiment 1. Raw performance scores and subjective ratings of overall workload from Experiment 1 are presented in Table 4. As shown in this table, reaction time of the basic Sternberg memory search task was prolonged as the difficulty was increased. This effect was observed under single task conditions (imposing demands upon specific processing stages) and under dual task conditions (low-bandwidth and high-bandwidth tracking with the memory task). Performing the memory search task with a low-bandwidth tracking task aggravated the RMS tracking errors. The tracking decrement was magnified when the memory task was made more difficult, particularly when the response load was imposed. RMS errors also increased in the high-bandwidth tracking task but reduced with the presence of a window. Each type of performance measure was standardized and combined to create the dependent variable to which the additive factors method was applied. The combined scores were then used to determine the locus of resources demanded by the easy tracking task and by the bandwidth manipulation.

-----  
Insert Table 4 about here  
-----

The primary purpose of localizing the resource demand is to substantiate the prior assumption of the degree of resource competition in different dual task conditions. It was assumed that the competition was relatively higher when a tracking task was performed concurrently with a response load memory task than when it was combined with other memory tasks.

Table 4 - Raw performance and subjective workload of each condition

Experiment 1

<u>Condition</u>	RMS	RT	AC	Workload
Basic Memory (BM)		.398	99.063	.983
Perceptual Load (PM)		.453	99.433	1.189
Central Load (CM)		.600	95.955	1.762
Response Load (RM)		.574	99.216	1.575
Low-bandwidth Tr (L)	.152			1.260
Low + Basic Mem (LB)	.165	.488	99.319	1.617
Low + Perceptual (LP)	.183	.514	99.102	1.804
Low + Central (LC)	.183	.665	94.539	2.314
Low + Response (LR)	.242	.594	97.788	2.076
High-bandwidth (HI)	.259			1.724
Narrow window (NA)	.225			2.490
Wide window (WI)	.220			2.211
High + Basic Mem (HB)	.272	.499	99.301	2.050
High + Perceptual (HP)	.293	.526	98.605	2.286
High + Central (HC)	.293	.661	94.715	2.812
High + Response (HR)	.365	.603	97.009	2.660

Through the additive factors method, this assumption can be tested. If the assumption is valid, then the tracking task should place greater demands on the stage of response selection and coordination. Time-sharing should be less efficient when the tracking task was executed concurrently with a response load memory task than when it was done with a perceptual or central load memory task. Similar logic is applied when examining the change in performance produced by an increase in tracking bandwidth.

Combined standardized scores were chosen to test the total resource competition in dual task conditions. By combining scores in this way, the measure was purposely made to be insensitive to any differential tradeoff in resource allocation between tasks in different conditions. The standardized scores of all the conditions are listed in the second column of Table 5.

-----  
 Insert Table 5 about here  
 -----

Figure 5 represents the interaction between the tracking difficulty and the Sternberg manipulation. The abscissa shows the effect of increasing memory search load by imposing demands upon central load (left panel) or upon response load (right panel). The solid line indicates the effect of these manipulations on single task Sternberg performance. The dashed line represents the dual task performance when the Sternberg manipulations were combined with a low-bandwidth tracking task. The dot-dashed line shows the dual task performance when these manipulations were combined with a high-bandwidth tracking task. Data presented in the top panels were generated by the weighting function  $0.5 * ZRMS + 0.25 * ( ZRT + ZAC )$  for dual task conditions and by the function  $0.50 * ( ZRT + ZAC )$  for all the single memory task conditions. This combination rule was chosen because subjects

Table 5 - Performance and subjective workload of each condition in comparable units - Experiment 1

<u>Condition</u>	<u>Standardized Score</u>		<u>Effect Size</u>		<u>Decrement Score</u>	
	P	S	P	S	P	S
Basic Memory (BM)	-.847	-1.192				
Perceptual Load (PM)	-.665	-.865	.88	1.47	.424	.333
Central Load (CM)	.527	-.134	2.50	2.01	2.634	1.265
Response Load (RM)	-.103	-.363	1.66	1.94	1.603	.961
Low-bandwidth Tr (L)	-1.021	-.814				
Low + Basic Mem (LB)	-.684	-.286	.41	1.63	.587	.834
Low + Perceptual (LP)	-.505	-.064	.60	1.47	.799	.940
Low + Central (LC)	.188	.465	.58	1.20	.947	1.303
Low + Response (LR)	.135	.225	.96	1.19	1.698	1.068
High-bandwidth (HI)	.258	-.208	1.87	.79	3.410	.753
Narrow window (NA)	-.151	.580	1.48	1.56	2.312	1.996
Wide window (WI)	-.209	.446	1.48	1.38	2.157	1.544
High + Basic Mem (HB)	-.012	.190	.47	1.43	.650	1.130
High + Perceptual (HP)	.228	.427	.70	1.35	.980	1.347
High + Central (HC)	.829	.904	.57	1.18	.924	1.735
High + Response (HR)	.957	.799	1.20	1.31	2.094	1.641

were instructed to give equal priority to each task and to respond as accurately and quickly as possible. Data depicted in the bottom panels were generated by the function ZRMS + ZRT + ZAC for dual task conditions and by the function ZRT + ZAC for all the single memory search tasks. The reason for this new combination format will be discussed latter.

-----  
Insert Figure 5 about here  
-----

It is noted that a good performance in an easy dual task condition could have resulted from the combination of a very low standardized RMS score with the memory task performance. Therefore, the large intercept difference between the dashed line (low-bandwidth tracking with a memory task) and dot-dashed line (high-bandwidth tracking with a memory task) was expected because the higher standardized RMS errors in the latter conditions. The effects of tracking baselines were removed in the numbers shown next to each point. These numbers equalled the combined increment from both the control tracking task and the basic memory search task. That is, difference was first obtained for each standardized dependent measure by subtracting the single task performance from the dual task performance. Difference standardized scores from all the dependent measures were then combined. These scores show the same relative relationships between the slopes as the relationships shown in the figure.

The resource demand of a low-bandwidth tracking task was tested by examining the interaction, as represented by the difference in the slope between the solid and dashed lines. If this tracking task imposes additional demands on a specific stage, performing it concurrently with a hard memory task which demands the same resources should elevate the

ORIGINAL PAGE IS  
OF POOR QUALITY

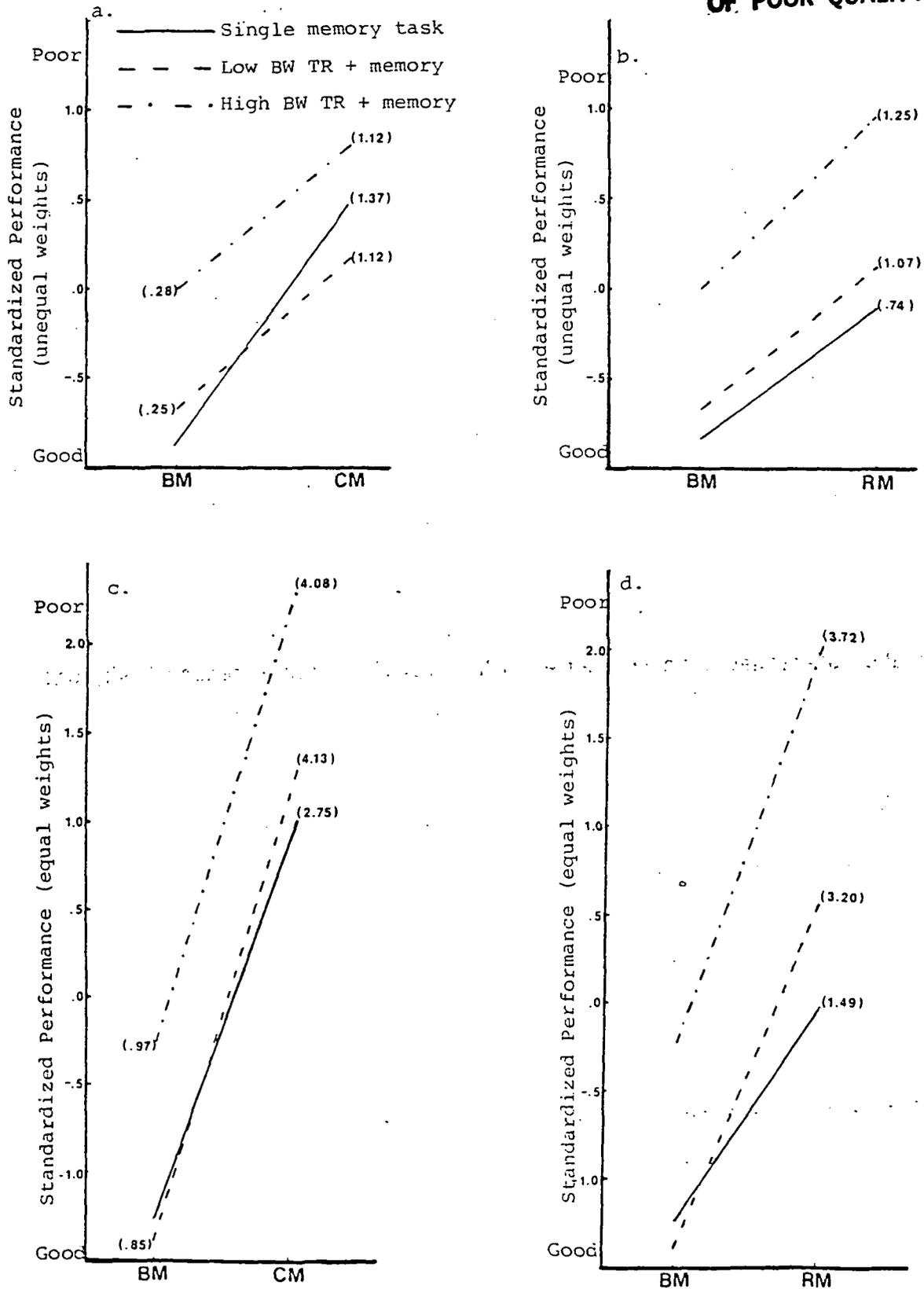


Figure 5. Interaction effects found via the additive factors method Experiment 1.

decrements to a greater extent than when it is performed with a basic memory search task. That is, the slope of the dashed line should be steeper than the slope of the solid line and the interaction between the two lines should be significant. The resource demand of increasing tracking bandwidth was examined by the same logic via the difference in the slope between the two types of dashed lines (i.e., the low and high bandwidth dual task conditions).

As shown in Figure 5a, performance of a central load memory task (CM) was worse than that of a basic memory search task (BM). However, the slope of single task conditions (the solid line) was steeper than that of dual task conditions (the dashed line). Following the additive factor logic, it appears that the low-bandwidth tracking task did not place extra demands on central processes. The combined dual task performance was actually better than single task performance in the central load condition and the interaction between the slopes was significant ( $F(1,14) = 18.34, p < .001$ ).

Results from dual task studies have shown that a tracking task demands resources for response selection and coordination (see Wickens, 1981 for a review). However, as shown in Figure 5b, the slopes were parallel between single memory task (the solid line) and dual task conditions (the dashed line). In other words, adding the low-bandwidth tracking task placed no extra demands on response resources.

The results shown in the top panels could be a consequence of the weights assigned to different performance measures. To test this possibility, equal weights were assigned to the three performance measures and the results are shown in the bottom panels. The interaction between the solid line (single tasks) and the dashed line (dual task conditions) in Figure 5c was not significant. This result, once again, indicated that

performing the low-bandwidth tracking task concurrently with the central load memory task did not enhance the demands for central resources. On the other hand, the effect of dual task loading was enhanced in the response load, relative to the central load condition (Figure 5d). The interaction as represented by the difference in the slope between the solid and dashed lines in Figure 5d was significant ( $F(1,14) = 11.23, p < .005$ ).

Neither the effect of dual task loading in cognitive processes nor the dual task effect in response load was enlarged further by increasing the tracking bandwidth. That is, increasing bandwidth did not produce any selective effect on the processing stages. The slopes of the dot-dashed lines in all the panels of Figure 5 were statistically parallel to those of the dashed lines.

Experiment 2. Raw performance and subjective workload data from Experiment 2 are shown in Table 6. In general, subjects were faster and more accurate when the processing code of a memory task was spatial than when it was verbal. The RMS errors were enlarged in both hard single task and dual task conditions. Employing a predictor display improved second-order tracking performance. But, the command display did not aid the tracking performance as much as anticipated.

-----  
Insert Table 6 about here  
-----

To test the prior assumption that resource competition was greater when a tracking task was performed with a spatial than with a verbal memory task, the additive factors method was again employed. If this assumption is right, then the dual task effect of performing a low-bandwidth tracking task in the spatial code condition should be greater than the effect in the

Table 6 - Raw performance and subjective workload of each condition

Experiment 2

<u>Condition</u>		RMS	RT	AC	Workload
Verbal Memory	(BM)		.623	94.686	1.320
Spatial Memory	(SM)		.557	97.270	0.790
Perceptual Load	(PM)		.576	96.619	0.948
Low-bandwidth Tr	(L )	.169			1.0
Low + verbal mem	(LV)	.209	.665	93.527	2.180
Low + spatial mem	(LS)	.206	.633	97.179	1.632
High-bandwidth	(HI)	.311			1.755
High + verbal mem	(HV)	.312	.673	94.504	3.118
High + spatial mem	(HS)	.318	.635	96.508	2.638
Second-order Tr	(SE)	.398			2.480
Predictor Display	(PD)	.252			2.108
Command Display	(CD)	.387			2.168
Second-order + PM		.442	.673	94.707	3.423
Predictive + PM		.330	.662	95.579	3.255
Command + PM		.449	.657	95.289	3.265

verbal code condition. The average standardized scores (i.e,  $0.5 * (ZRT + ZAC)$  for all the single memory task conditions and  $0.5 * ZRMS + 0.25 * (ZRT + ZAC)$  for dual task conditions) were analyzed to test the total cost in a dual task condition, independent of the tradeoff in resource allocation. The standardized scores of all the conditions are presented in the second column of Table 7 and graphically depicted in Figure 6.

-----  
 Insert Table 7 & Figure 6 about here  
 -----

As shown in Figure 6a, single task performance of a spatial code condition (SM) was better than that of the verbal code baseline memory condition (BM: the solid line). The advantage of performing the spatial memory task also held when it was combined with a low-bandwidth tracking task (the dashed line). However, the superiority in the dual task condition was relatively smaller than that in the single task condition and the slope of the dashed line is less steep than that of the solid line. Decrements occurred when the low-bandwidth tracking task was added to the spatial memory task. On the other hand, performance did not deteriorate by performing the tracking task concurrently with the verbal memory task. As predicted by the multiple resources model, the effect of dual task loading was enlarged in the spatial code condition, relative to the verbal code condition. The difference in the slope between the solid line (single tasks) and the dashed line (dual task conditions) was significant ( $F(1,14) = 5.83, p < .05$ ). Increasing bandwidth did not place additional demands on a specific processing code, as indicated by the result that the interaction between the two types of dashed lines was not significant.

Table 7 - Performance and subjective workload of each condition in comparable units - Experiment 2

<u>Condition</u>	<u>Standardized Score</u>		<u>Effect Size</u>		<u>Decrement Score</u>	
	P	S	P	S	P	S
Verbal Memory (BM)	.052	-.682				
Spatial Memory (SM)	-.672	-1.442				
Perceptual Load (PM)	-.474	-1.100				
Low-bandwidth Tr (L)	-1.271	-.829				
Low + verbal mem (LV)	-.218	.239	.697	2.065	.723	.907
Low + spatial mem (LS)	-.568	-.176	.729	1.996	.772	.655
High-bandwidth (HI)	-.031	-.053	2.755	2.605	3.053	.671
High + verbal mem (HV)	.171	.731	.193	1.633	.270	1.460
High + spatial mem (HS)	-.050	.543	.406	2.017	.513	1.215
Second-order Tr (SE)	.688	.384	3.403	2.575	4.907	1.316
Predictor Display (PD)	-.527	.026	1.298	1.450	1.775	.986
Command Display (CD)	.599	.026	3.821	1.276	4.677	1.039
Second-order + PM	.695	.848	.674	1.686	1.095	1.520
Predictive + PM	.138	.754	.773	1.725	1.335	1.536
Command + PM	.643	.719	.763	1.601	1.166	1.518

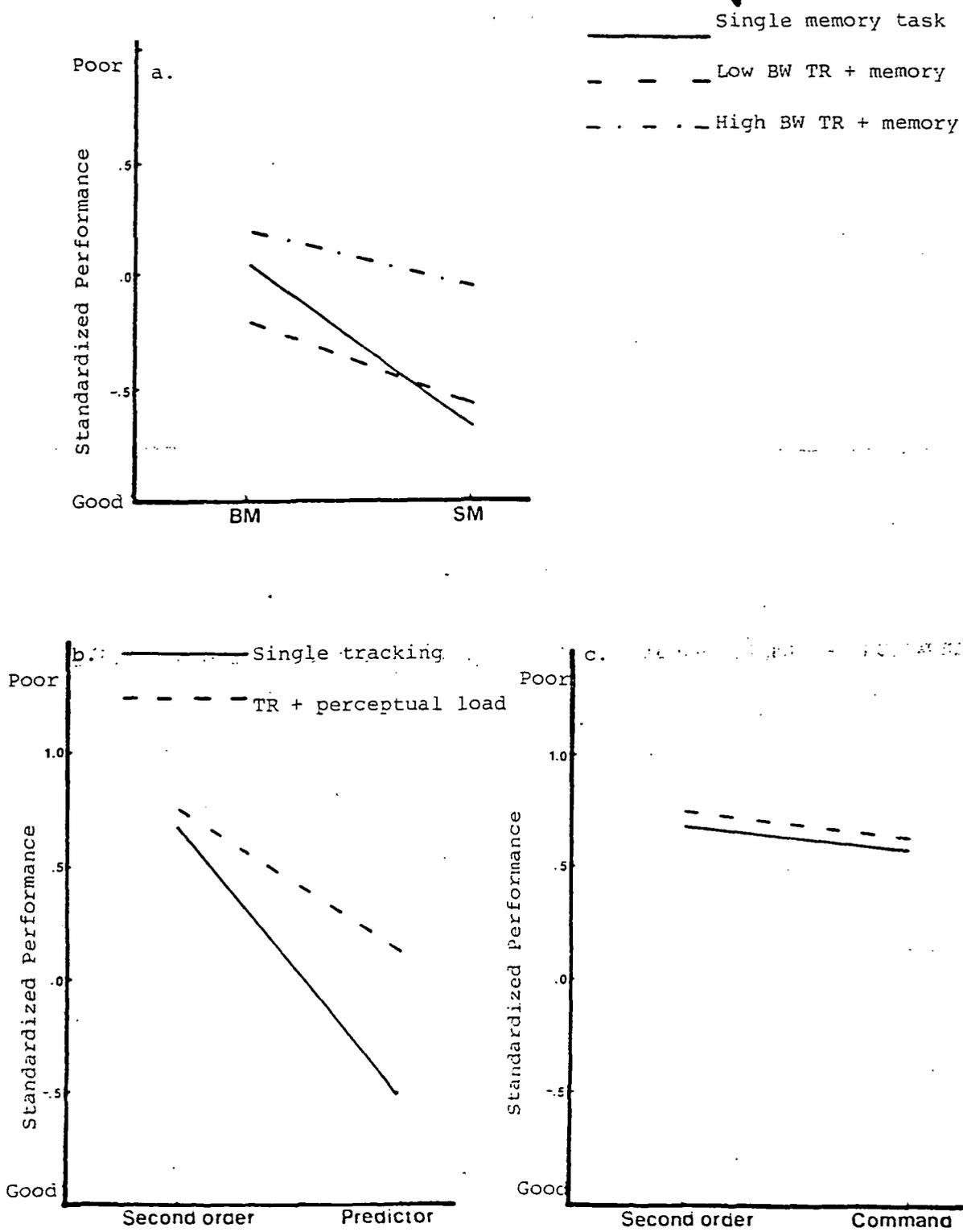


Figure 6. Interaction effects found via the additive factors method  
Experiment 2

The additive factors method was also applied to test the prior assumption of the effect of a predictor display. It was presumed that this display demanded relatively more perceptual resources than did an unaided display because two distinct objects were processed to determine control inputs. Data, supporting this assumption, are shown in Figure 6b. The solid line represents single tracking task performance and the dashed line shows the performance when the tracking task was added to a perceptual load memory search task.

The benefit of tracking with a predictor display was pronounced in both single and dual task conditions, but the advantage was relatively larger in the former than in the latter condition. The slope of the solid line (single tasks) was much steeper than the slope of the dashed line (dual task conditions). That is, adding a perceptual load memory task had a small effect on the unaided tracking task. But, performance dropped to a great degree when the "predictor display" tracking task was performed concurrently with the perceptual load memory task. The interaction between the two lines in Figure 6b was significant ( $F(1,14) = 61.75, p < .0001$ ).

The perceptual demand of the command display was also tested by the additive factors method. The solid line in Figure 6c represents single task performance and the dashed line indicates the effect of adding the perceptual load memory task to the tracking task. There was no main effect of the display type and thus no benefit from the command display. In addition, the two lines in Figure 6c are parallel and there was no interaction. That is, the command display imposed no greater demands on perceptual resources than did the unaugmented display.

In summary, the results of the additive factors analyses from Experiments 1 and 2 indicate that a tracking task (independent of bandwidth)

is more spatial than verbal in its processing code and these demands are more localized in response processes than in central memory processes. The locus of resource demand of increasing bandwidth appears to be stage-non-specific. This evidence confirms the previous tentative finding in Wickens and Derrick's study (1981).

Experiment 3. The Raw data from Experiment 3 are shown in Table 8.

Tracking performance became worse when a basic dual-axis tracking task (BT - one cursor controlled by one joystick) was combined with any of the memory tasks. Performance also decreased when the cursor was controlled by two joysticks (CT) and dropped further when the display was separated (DT - two cursors tracked by one stick). Reaction times and errors of a basic memory task generally increased when a tracking task was added to the memory task, whereas RT was also lengthened from the basic memory task by increasing memory load, employing spatial stimuli, and using a speech response.

-----  
Insert Table 8 about here  
-----

The standardized performance measures were weighted and combined in the same format as that used in the second experiment. The standardized scores of all the conditions are shown in the second column of Table 9. The additive factors method was then applied to localize the processing demands of dual-axis tracking tasks in various display/control formats. Determining the locus of resources demanded by these dual-axis tracking tasks was particularly important because there had been no previous evidence to support the suppositions that (1) separating an integrated cursor into two distinct error indicators imposes more demands upon perceptual/cognitive

Table 8 - Raw performance and subjective measure of each condition

Experiment 3

<u>Condition</u>	RMS	RT	AC	Workload
Basic Memory (BM)		.482	97.185	.891
Central Load (CM)		.590	93.413	2.10
Spatial Memory (SM)		.543	96.526	1.018
Speech (SP)		1.066	97.768	.661
Basic Dual-Axes (BT)	.153			1.000
Separated display (DT)	.229			1.630
Separated control (CT)	.181			1.546
BT + basic mem (BB)	.171	.543	96.746	1.652
BT + central load (BC)	.186	.650	92.289	3.237
BT + spatial mem (BS)	.191	.640	95.821	1.888
BT + speech (BP)	.176	1.098	96.573	1.596
DT + basic mem (DB)	.263	.570	96.045	2.380
DT + central load (DC)	.275	.661	93.068	4.360
DT + spatial mem (DS)	.280	.654	94.206	2.716
DT + speech (DP)	.270	1.124	94.823	2.239
CT + basic mem (CB)	.186	.549	96.713	2.356
CT + central load (CC)	.196	.652	91.923	4.098
CT + spatial mem (CS)	.196	.654	95.845	2.730
CT + speech (CP)	.195	1.125	95.999	2.118

resources, and (2) separating control increases the demands on response-related resources.

-----  
 Insert Table 9 about here  
 -----

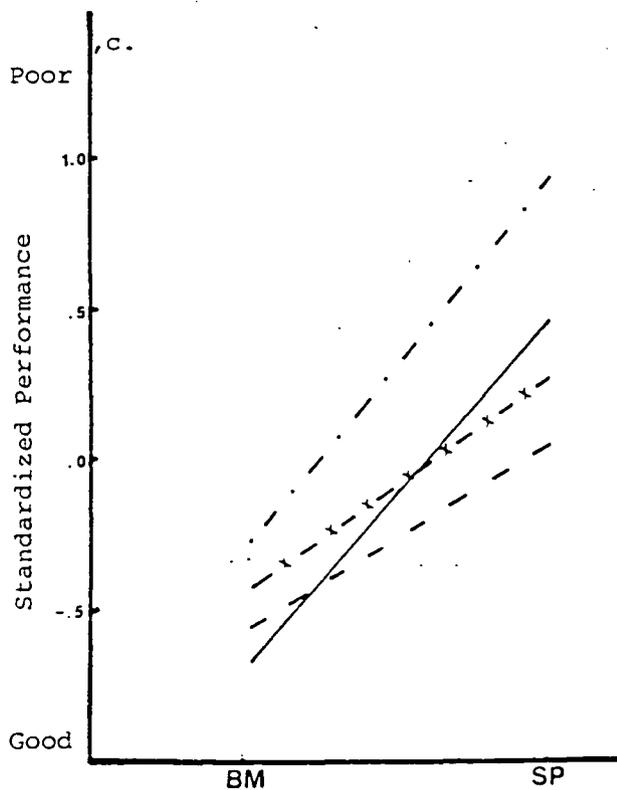
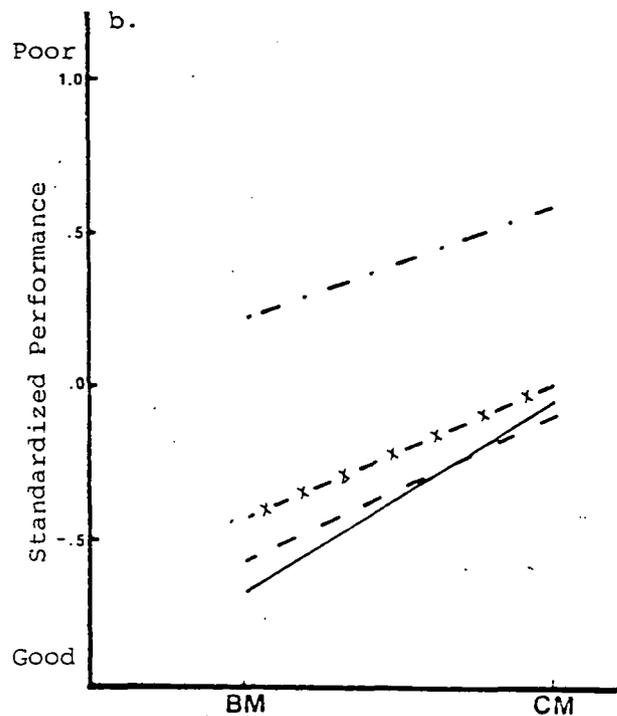
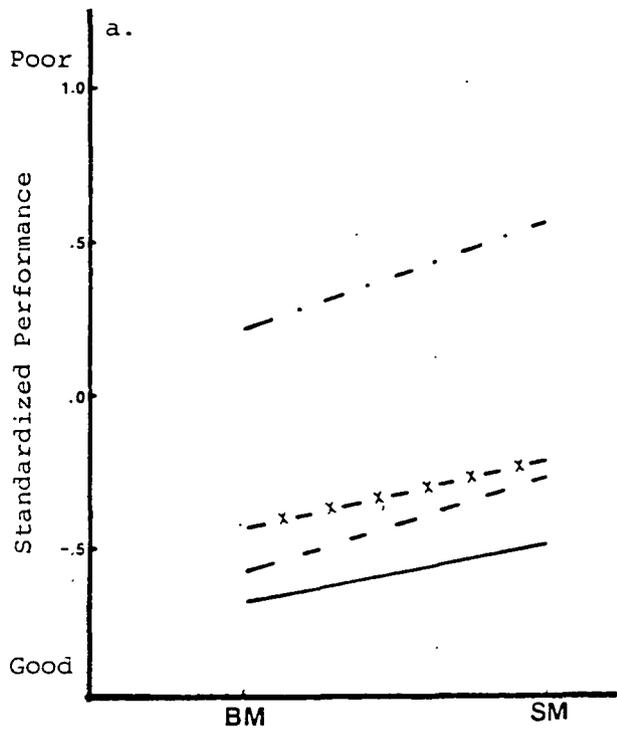
If the assumptions are valid, then (1) the effect of display separation and the effect of loading central processes of the Sternberg memory search task should be interactive, and (2) control separation and the basic memory search task (manual responses were required) should compete for response-related resources. The relevant data from Experiment 3 are displayed in Figure 7. The abscissa in panels a, b, and c indicates the effect of changing a processing characteristic of the basic verbal Sternberg memory search task (BM) by the processing code (SM), central load (CM), and response mode (manual to speech - SP) respectively. The solid lines represent the performance under single memory task conditions. The dashed lines show the performance when a basic dual-axis tracking task (BT) was added to each of the memory tasks. The dot-dashed lines indicate the performance of doing one of the memory tasks concurrently with a tracking task whose display was separated (DT). The x-dashed lines illustrate the performance when the memory task was combined with a tracking task in which the control was separated (CT).

-----  
 Insert Figure 7 about here  
 -----

The dual task effect of performing the integrated dual-axis tracking task was not enhanced in the spatial code or central load condition, as compared with the basic verbal memory task. Furthermore, separating display

Table 9 - Performance and subjective measure of each condition in comparable units - Experiment 3

<u>Condition</u>	<u>Standardized Score</u>		<u>Effect Size</u>		<u>Decrement Score</u>	
	P	S	P	S	P	S
Basic Memory (BM)	-.687	-1.453				
Central Load (CM)	-.060	.015	1.172	1.776	1.543	1.627
Spatial Memory (SM)	-.489	-.968	.568	.622	.685	.372
Speech (SP)	.459	-1.616	2.924	-.216	5.641	-.043
Basic Dual-Axes (BT)	-.925	-.756				
Separated display (DT)	.302	-.118	1.873	2.205	3.680	.730
Separated control (CT)	-.476	-.231	.764	1.518	1.346	.634
BT + basic mem (BB)	-.569	-.084	.563	2.829	.781	.931
BT + central load (BC)	-.103	.750	.662	1.897	1.169	1.956
BT + spatial mem (BS)	-.262	.103	.885	2.553	1.439	1.019
BT + speech (BP)	.053	-.130	.526	2.478	.802	.888
DT + basic mem (DB)	.235	.404	.709	2.147	1.331	1.411
DT + central load (DC)	.586	1.198	.692	1.788	1.490	2.895
DT + spatial mem (DS)	.560	.584	.972	1.821	1.938	1.615
DT + speech (DP)	.930	.342	.699	2.027	1.470	1.269
CT + basic mem (CB)	-.442	.389	.342	2.147	.494	1.431
CT + central load (CC)	-.003	1.090	.428	1.662	.763	2.638
CT + spatial mem (CS)	-.210	.571	.552	1.739	.952	1.679
CT + speech (CP)	.266	.230	.416	1.757	.758	1.177



- Single memory task
- - - Basic dual-axis + Mem
- . - . Separated display + Mem
- x - x - x Separated control + Mem

Figure 7. Interaction effects found via the additive factors method Experiment 3

or control placed no greater demands on the spatial code or the central processes than did the basic dual-axis tracking task. In other words, the four lines are statistically parallel in Figure 7a and in Figure 7b.

When the memory task was responded to with speech, a concurrent benefit was found in some of the dual task conditions. The slope of the dashed line (integrated dual-axis tracking with a memory task) in Figure 7c is less steep than that of the solid line (single memory tasks) and the interaction between the two lines was significant ( $F(1,13) = 31.65, p < .0001$ ). The dashed line did not interact with the dot-dashed line (separated-display tracking with a memory task) or with the x-dashed line (separated-control tracking with a memory task). Thus, it appears that neither the display nor the control separation demanded a greater load from manual resources than did the integrated tracking task.

#### Rating reliability

Correlations between the ratings collected in Sessions 3 and 4 were computed across all the subjects and the conditions for each rating scale. Table 10 shows these correlations for both the ratings recorded at the end of each trial and the ratings recorded at the end of a session. Although both post trial and post session reliabilities were fairly high, the latter ratings were more reliable than the former. For Experiments 1, 2, and 3 respectively, the reliabilities of post trial ratings were 0.67, 0.69, and 0.70. The reliability of post session ratings in all the three experiments was 0.83.

-----  
Insert Table 10 about here  
-----

Table 10 - Rating reliability

	Experiment 1		Experiment 2		Experiment 3	
	Post trial, Post session		Post trial, session		Post trial, session	
Overall workload	.72	.79	.69	.84	.67	.88
Perceptual effort	.67	.75	.70	.86	.72	.90
Mental effort	.73	.83				
Response effort	.69	.85	.63	.85	.77	.83
Time pressure	.53	.85	.71	.79	.68	.79
Stress	.66	.87	.68	.88	.52	.79
Performance	.78	.76				
Overall effort	.57	.84				
Excess capacity	.68	.90				
Overall	.67	.83	.69	.83	.70	.83
Task complexity			.68	.73	.70	.71
Capacity used			.77	.84	.80	.90
Cooper/Harper			.70	.85	.78	.82

### Predicting global workload from specific rating scales

Table 11 shows the correlations between the overall workload ratings and ratings on every other scale used in the experiment. As shown in this table, most ratings were highly correlated with the ratings on the overall workload scale. A multiple regression analysis was conducted to determine which variables significantly predict the ratings on the overall workload scale.

-----  
Insert Table 11 about here  
-----

The BMDP 9R program was used to execute the multiple regression analysis. The program was written to "estimate regression equation for "best" subsets of predictor variables" (BMDP, 1981, p. 264). In practice, this program selects a subset of predictors that account for most of the variance of the dependent variable. This program starts with a subset of any one predictor. For each set, the program then selects the next predictor that significantly maximizes the adjusted R square. The adjusted R square indicates the amount of variance of the dependent variable accounted for by the new subset. The program stops selection when adding a new predictor to the subset does not significantly improve the adjusted R square. The program executes this analysis for all possible subsets and selects the best subset for predicting the dependent variable. In addition, the program also lists the T-value corresponding to the regression weight of each predictor in a subset. Therefore, the significance of the contribution of each predictor can be determined.

The results of the analysis are presented in Table 12. In Experiment 1, the combination of four predictors accounted for most of the variance in

Table 11 - Correlations between overall workload ratings and other ratings

Scale	Experiment 1	Scale	Experiment 2	Experiment 3
PE	.68	P/ME	.86	.90
ME	.81	RE	.78	.46
RE	.67	TI	.80	.54
TI	.65	ST	.73	.30
ST	.74	CM	.73	.61
PR	.18	CA	.57	.39
OE	.68	MCH	.82	.58
EX	-.68			

PE - perceptual effort, ME - mental effort, RE - response effort

TI - time pressure , ST - stress level , PR - performance

OE - overall effort , EX - excessive capacity

P/ME - perceptual/mental effort, CM - task complexity

CA - capacity used , MCH - modified Cooper/Harper

the overall workload ratings (90%). The other predictors did not significantly improve the adjusted R square. These four predictors were perceptual effort (PE), mental effort (ME), time-demand (TI), and stress (ST). Among the four predictors, mental effort had the highest regression weight and accounted for more variance than the other three predictors. In Experiment 2, workload was best predicted by perceptual/mental effort (P/ME), time-demand (TI), capacity used (CA), and the modified Cooper/Harper scale (MCH). These variables accounted for 95% of the variance in the overall workload ratings. Ratings on P/ME had a higher regression weight than the other three predictors. In Experiment 3, 93% of the variance in the overall workload ratings was accounted for by the combination of perceptual/mental effort (P/ME), response effort (RE), stress (ST), and task complexity (CM). The P/ME scale had the highest regression weight among the four predictors. Across the three experiments, perceptual/mental effort scale was the one that consistently predicted the ratings on the overall workload scale.

-----  
Insert Table 12 about here  
-----

#### The perceived structure of subjective difficulty

Similarity judgment data were analyzed by the SINDSCAL program. The interpretation of the disclosed dimensions was based upon (1) the correlations with the unidimensional ratings, and (2) task location on each of the dimensions.

As shown in Table 13, most unidimensional ratings were correlated with subjective dimensions 1, 3, and 4 in the data of Experiment 1, whereas

Table 12 - Predicting Overall workload from specific scales

Experiment 1

Adjusted R square = 0.90

$$WK = .095 + .263 PE + .319 ME + .154 TI + .292 ST$$

Experiment 2

Adjusted R square = 0.90

$$WK = .03 + .549 P/ME + .359 TI - .10 CA + .13 MCH$$

Experiment 3:

Adjusted R square = 0.93

$$WK = .085 + .979 P/ME - .085 RE - .317 ST + .295 CM$$

response effort ratings were related to all of the four dimensions. Data of Experiment 2 (Table 14) showed that the two subjective dimensions were correlated with the following scales: overall workload, perceptual/mental effort, response effort, task complexity, and the modified Cooper/Harper scale. The three subjective dimensions disclosed from the data of Experiment 3 (Table 15) were correlated with overall workload, task complexity, capacity used, and the modified Cooper/Harper scale. Overall workload ratings were related to most of the subjective dimensions across different task configurations.

-----  
Insert Tables 13, 14, & 15 about here  
-----

Replicating Yeh and Wickens' result (1984), a dimension of processing codes was revealed from the SINDSCAL solution of Experiment 1. Figure 8 presents two of the subjective dimensions from Experiment 1. As shown in this figure, all the single memory tasks demanding verbal codes had positive weights on Dimension 2. The two single tracking tasks (low-bandwidth tracking - L, and high-bandwidth tracking - H) had negative weights. Dual tasks that demanded both verbal and spatial codes were located between the two clusters.

-----  
Insert Figure 8 about here  
-----

Two of the subjective dimensions disclosed from the data of Experiment 1 were associated with the demand on perceptual/central resources. As shown in Figure 8, the first dimension was related to the processing stage. Basic memory tasks and perceptual load memory tasks had positive weights on this

Table 13 - Correlations of SINDSCAL dimension weights with  
unidimensional ratings (Experiment 1)

Unidimensional scale	Dimension			
	1	2	3	4
Overall workload	-.75	-.43	-.87	.64
Perceptual effort	-.48	-.41	-.89	.59
Mental effort	-.80	-.32	-.85	.58
Response effort	-.68	-.57	-.87	.64
Time pressure	-.71	-.53	-.87	.63
Stress	-.77	-.43	-.86	.64
Performance	.71	.44	.80	-.74
Overall effort	-.71	-.52	-.86	.68
Excess capacity	.73	.46	.89	-.62

Note. The following critical values can be used to evaluate the  
correlations reported in the table:  $r(12) = .532$ ,  $p < .05$ ;  
 $r(12) = .612$ ,  $p < .01$

Table 14 - Correlations of SINDSCAL dimension weights with  
 unidimensional ratings (Experiment 2)

Unidimensional scale	Dimension	
	1	2
Overall workload	.70	.76
Perceptual/Mental effort	.61	.75
Response Effort	.72	.79
Time pressure	.72	.76
Stress	.50	.74
Task complexity	.67	.84
Capacity used	.57	.74
Cooper/Harper	.67	.79

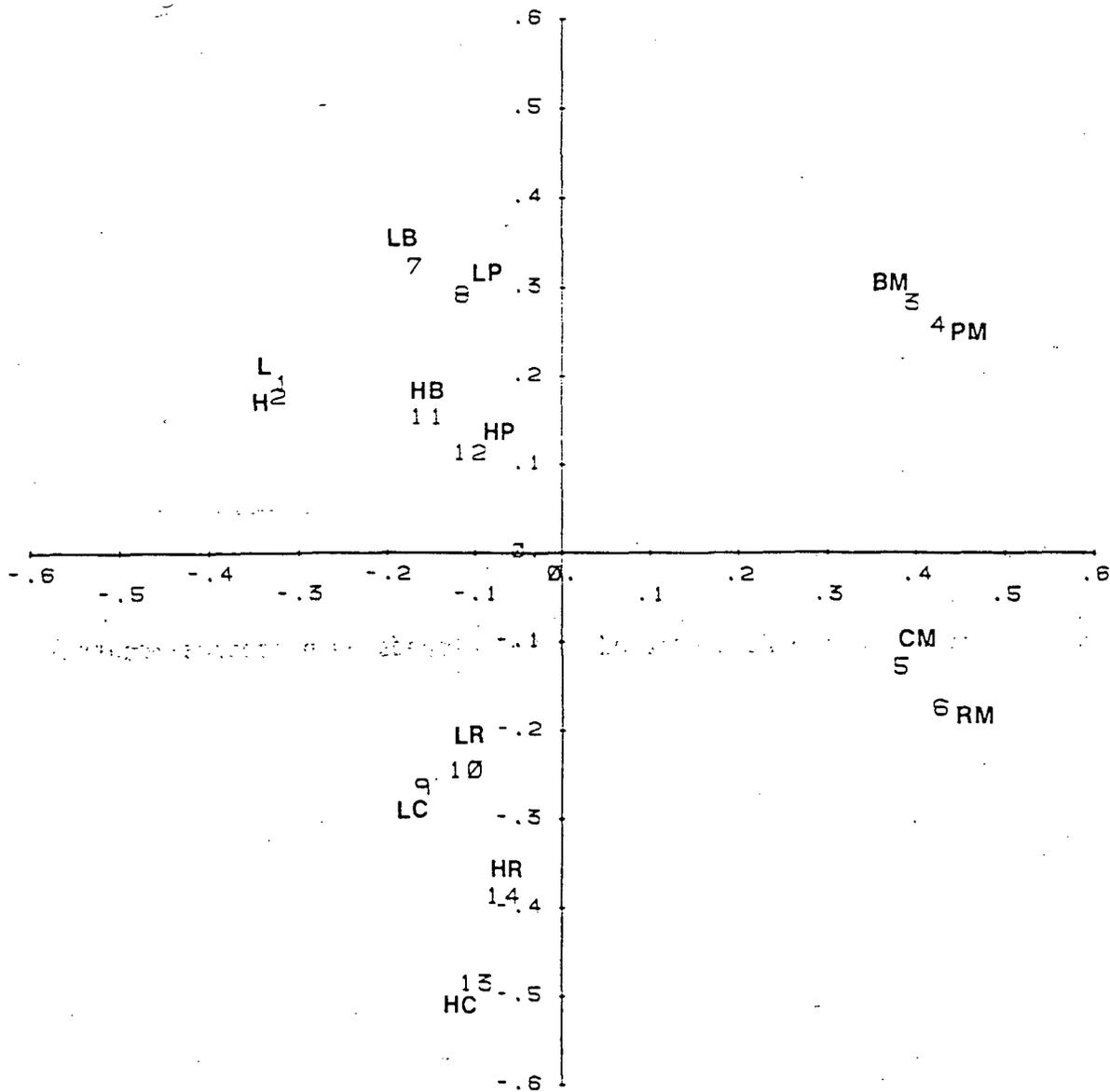
Note. The following critical values can be used to evaluate the  
 correlations reported in the table:  $r(9) = .605$ ,  $p < .05$ ;  
 $r(9) = .685$ ,  $p < .01$

Table 15 - Correlations of SINDSCAL dimension weights with  
 unidimensional ratings (Experiment 3)

Unidimensional scale	Dimension		
	1	2	3
Overall workload	.52	.82	-.77
Perceptual/Mental effort	.39	.77	-.81
Response Effort	.46	.90	-.70
Time pressure	.18	.73	-.84
Stress	-.00	.21	-.68
Task complexity	.51	.79	-.79
Capacity used	.55	.79	-.79
Cooper/Harper	.52	.81	-.78

Note. The following critical values can be used to evaluate the  
 correlations reported in the table:  $r(14) = .497$ ,  $p < .05$ ;  
 $r(14) = .623$ ,  $p < .01$

ORIGINAL PAGE IS  
OF POOR QUALITY



HORIZONTAL AXIS: DIMENSION II  
VERTICAL AXIS : DIMENSION I

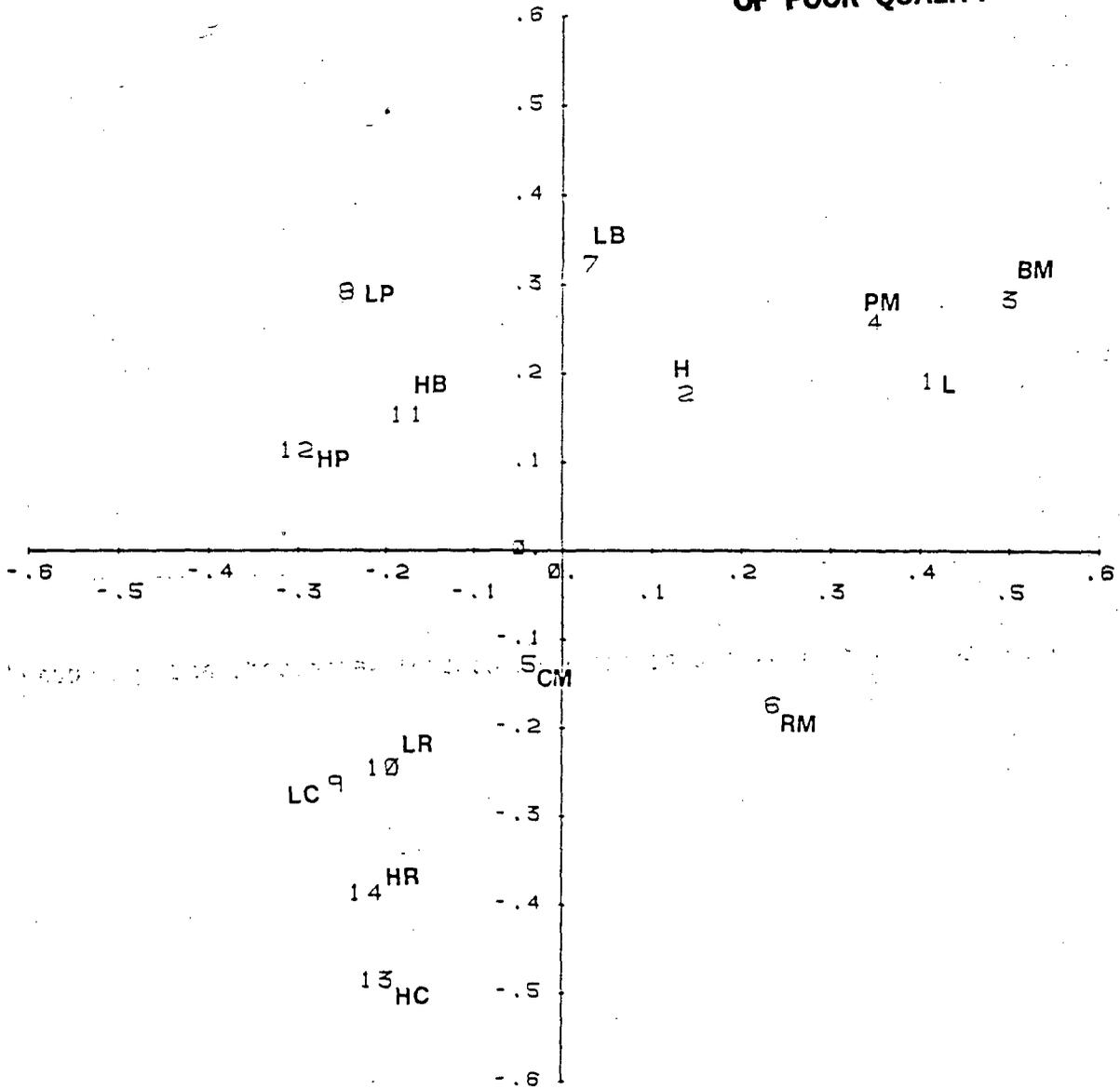
Figure 2. SINDSCAL solution from the similarity ratings of task difficulty - Experiment 1

dimension. Central load and response load memory tasks had negative weights. Figure 9 represents Dimensions 1 and 3 from the SINDSCAL solution of Experiment 1. Dual tasks that demanded more resources to coordinate (all except the condition in which a low-bandwidth tracking was combined with a basic memory task - LB) had negative weights on Dimension 3. The central load memory task also had negative weights on this dimension. The other tasks demanding less cognitive resources had positive weights on Dimension 3.

-----  
Insert Figure 9 about here  
-----

All the three dimensions disclosed from the SINDSCAL solution in the data of Experiment 3 were related to the demands on perceptual/central resources. Dimension 1 in Figure 10 was tied to the demands for working memory. All the single memory tasks, which required that information be held in working memory, had negative weights. On the other hand, single tracking tasks which did not demand memory capacity were located on the top of the dimension. Dual tasks, combining the two types of tasks, were located between the two extremes. Dimension 2 (Figure 11) also appeared to be associated with the demands for working memory. All the single tasks had negative weights whereas dual task conditions which required executive management had positive weights on Dimension 2. On Dimension 3 (Figure 10), tasks that demanded more perceptual/central resources had negative weights. These tasks included central load memory task, tasks in which subjects perceived two cursors, and dual tasks in which a tracking task competed with a spatial memory task. The other tasks had positive weights on Dimension 3.

ORIGINAL PAGE IS  
OF POOR QUALITY



HORIZONTAL AXIS: DIMENSION III  
VERTICAL AXIS: DIMENSION I

Figure 9. SINDSCAL solution from the similarity ratings of task difficulty - Experiment 1

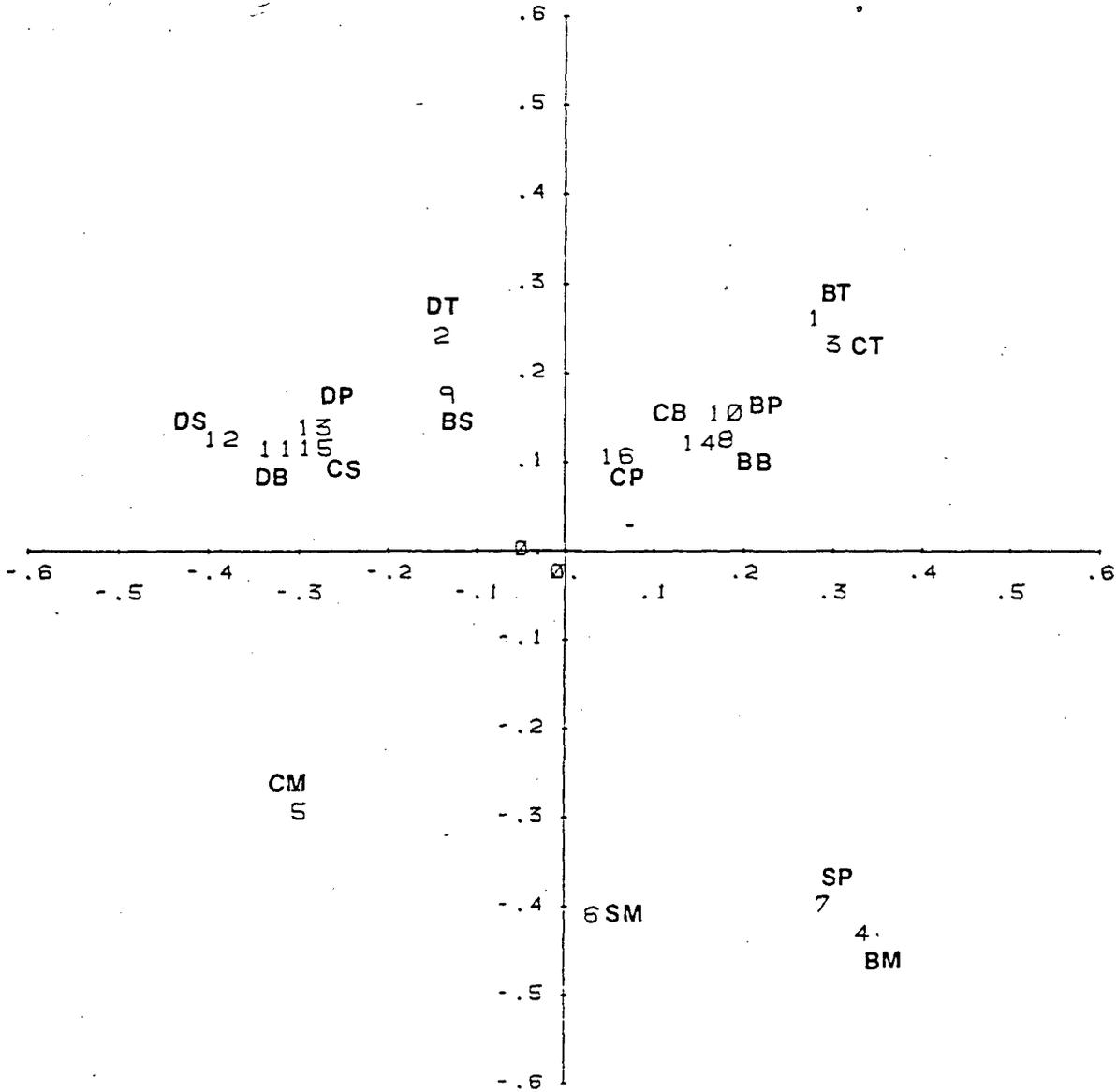
-----  
Insert Figures 10 & 11 about here  
-----

### Dissociation between performance and subjective measures

Comparable performance and subjective ratings of overall workload from each condition are presented in Tables 5, 7, and 9 for the three experiments respectively. The tasks that subjects performed in each experiment are listed in the first column and the codes in parentheses are the abbreviations of each condition. In dual task conditions, the first letter indicates the type of the tracking task (L - low bandwidth, H - high bandwidth, B - basic dual-axis, D - dual-axis tracking with display separated, and C - dual-axis tracking with control separated) and the second letter represents the type of the memory task (B - basic verbal, P - perceptual load, C - central load, R - response load, S - spatial code, and P - speech mode). Comparable performance and subjective workload measures, derived from the three types of analysis techniques for testing dissociation, are listed to the right. Since both the effect sizes and decrement scores are computed as the changes from the baseline condition(s), scores for the baseline conditions (easy tracking and basic memory search task) are both zero.

Tables 16 - 19 list the results of the ANOVA's testing of each dissociation predicted by the proposed theory. In these tables, comparisons relevant to the tested dissociation are organized in the first column by the experiment in which each comparison was made. Following each comparison, the F and P values of the interaction effect in the analysis of each of the three types of measures are listed.

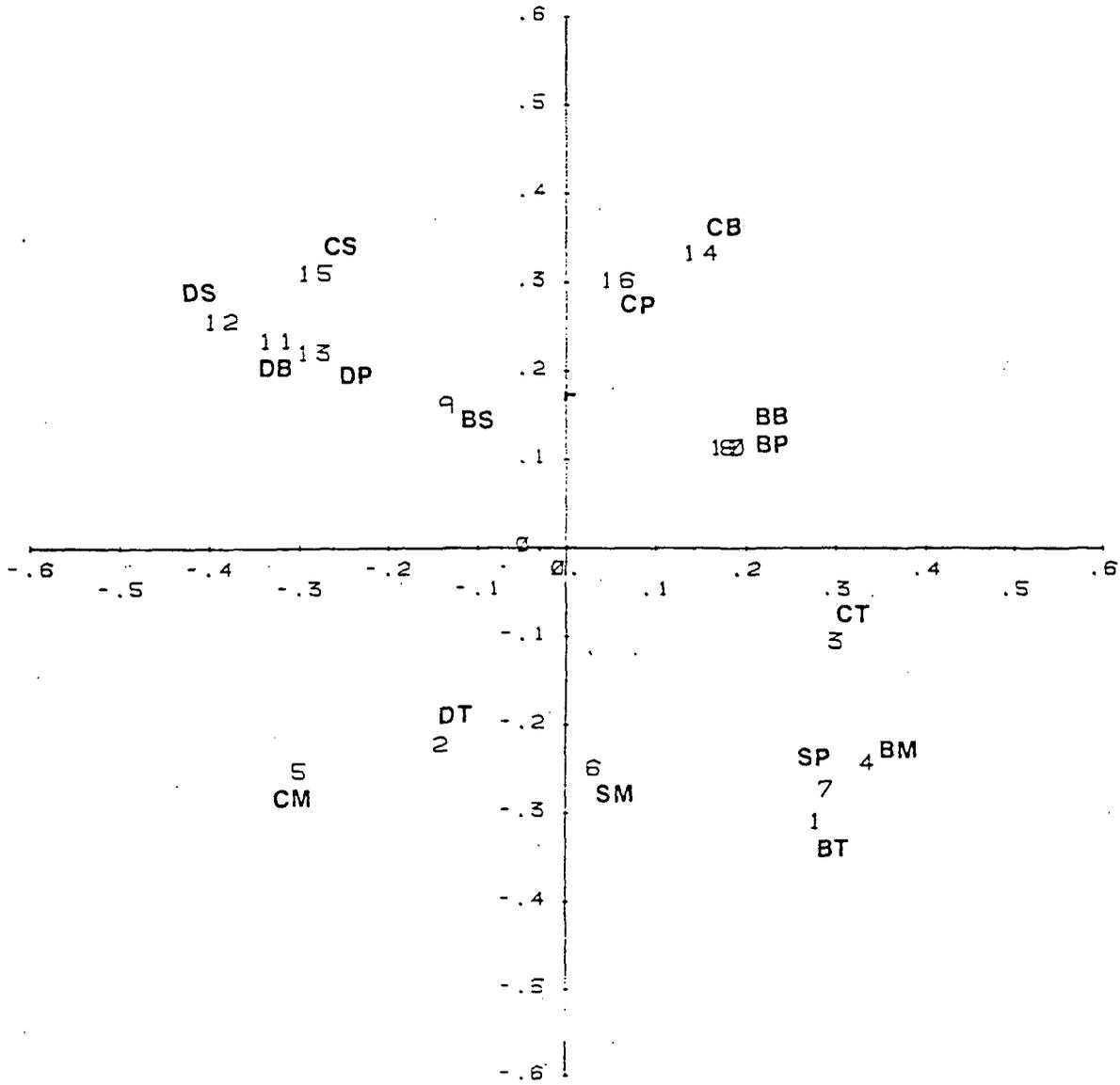
ORIGINAL PAGE IS  
OF POOR QUALITY



HORIZONTAL AXIS: DIMENSION III  
VERTICAL AXIS : DIMENSION I

Figure 10. SINDSCAL solution from the similarity ratings of task difficulty - Experiment 3

ORIGINAL PAGE IS  
OF POOR QUALITY



HORIZONTAL AXIS: DIMENSION III  
VERTICAL AXIS : DIMENSION II

Figure 11. SINDICAL solution from the similarity ratings of task difficulty - Experiment 2

The results listed in these tables are also represented graphically in Figures 12 - 17. The Y-value of the ending point of each vector in these figures represents the amount of performance decrements produced by a specific manipulation. In a single memory task condition, this amount was derived by averaging the normalized latency and accuracy decrement scores. In a dual task condition, it was combined from three normalized decrement scores ( $0.5 * \text{RMS decrements} + 0.25 * (\text{RT decrements} + \text{AC decrements})$ ). This combined decrement score estimates the total cost in a dual task condition, independent of the tradeoff in resource allocation or between speed and accuracy. The X-value of the same point shows the amount of subjective workload increased from the control condition(s), as assessed by the overall workload scale. Thus, the angle of the vector indicates its relative influence on performance vs. subjective ratings, whereas the length represents the total "strength" of a manipulation.

Data will not be discussed figure by figure, but rather in the context of each source of dissociation.

Perceptual vs. Response resources. According to the proposed theory, the vector for a response load manipulation should point relatively more toward the performance axis than the vector for a perceptual or central load manipulation. All of the relevant results from the three experiments are presented in Table 16.

-----  
 Insert Table 16 about here  
 -----

Vectors a, b and c in Figure 12 are three conditions relevant to this dissociation in Experiment 1. It was anticipated that both the perceptual

Table 16 - Results of the interaction effect in ANOVA tests

Contrast: Perceptual vs. Response resources.

<u>Comparison</u>	<u>Standardized Score</u>		<u>Effect Size</u>		<u>Decrement Score</u>	
	F value	P <	F value	P <	F value	P <

Experiment 1

PM vs. RM	0.16	-	3.21	.10	5.84	.03
-----------	------	---	------	-----	------	-----

CM vs. RM	4.34	.06	0.26	-	5.64	.04
-----------	------	-----	------	---	------	-----

Experiment 2

SE vs. HI	2.92	.11	0.85	-	10.83	.006
-----------	------	-----	------	---	-------	------

Experiment 3

DT vs. CT	22.67	.0004	0.494	-	59.29	.00001
-----------	-------	-------	-------	---	-------	--------

The effect of the number of display elements

SE vs. PD	22.28	.0003	2.30	.20	65.17	.00001
-----------	-------	-------	------	-----	-------	--------

CD vs. PD	52.53	.00001	16.78	.0001	79.85	.00001
-----------	-------	--------	-------	-------	-------	--------

SE vs. CD	2.00	.20	5.33	.025	0.03	-
-----------	------	-----	------	------	------	---

load (Vector a) and central load condition (Vector b) would be further from the performance axis than the response load condition (Vector c). In the case of perceptual load, this was true and the dissociation (PM vs. RM in Table 16) was significant in the analysis of decrement scores ( $F(1,13) = 5.84, p < .03$ ). However, Vector b (central load) points to the left and above Vector c (response load), in contrast to the prediction. Performance of the central load memory task was much worse than what its subjective workload would indicate. Statistics of two measures (standardized scores and decrement scores) supported this reverse dissociation (CM vs. RM in Table 16).

-----  
Insert Figure 12 about here  
-----

The processing stage dissociation in the data of Experiment 2 is shown in Figure 13. The vector for the second-order tracking task (demands stage-related resources) was predicted to point relatively more toward the subjective workload axis than the vector for the high-bandwidth tracking task (demands stage-non-specific resources). However, the difference in performance between the high-bandwidth tracking task (Vector a) and the second-order tracking task (Vector b) is greater than the difference between their subjective workload ratings. That is, performance decrements of the second-order tracking task were underrated by the subjective workload, relative to the high-bandwidth tracking task. This dissociation contradicted the prediction and the earlier finding of Wickens and Yeh's study (1983). The dissociation was significant in the analysis of decrement scores ( $F(1,14) = 10.83, p < .006$ ) and close to the significance level by testing the standardized scores ( $p < .11$ ).

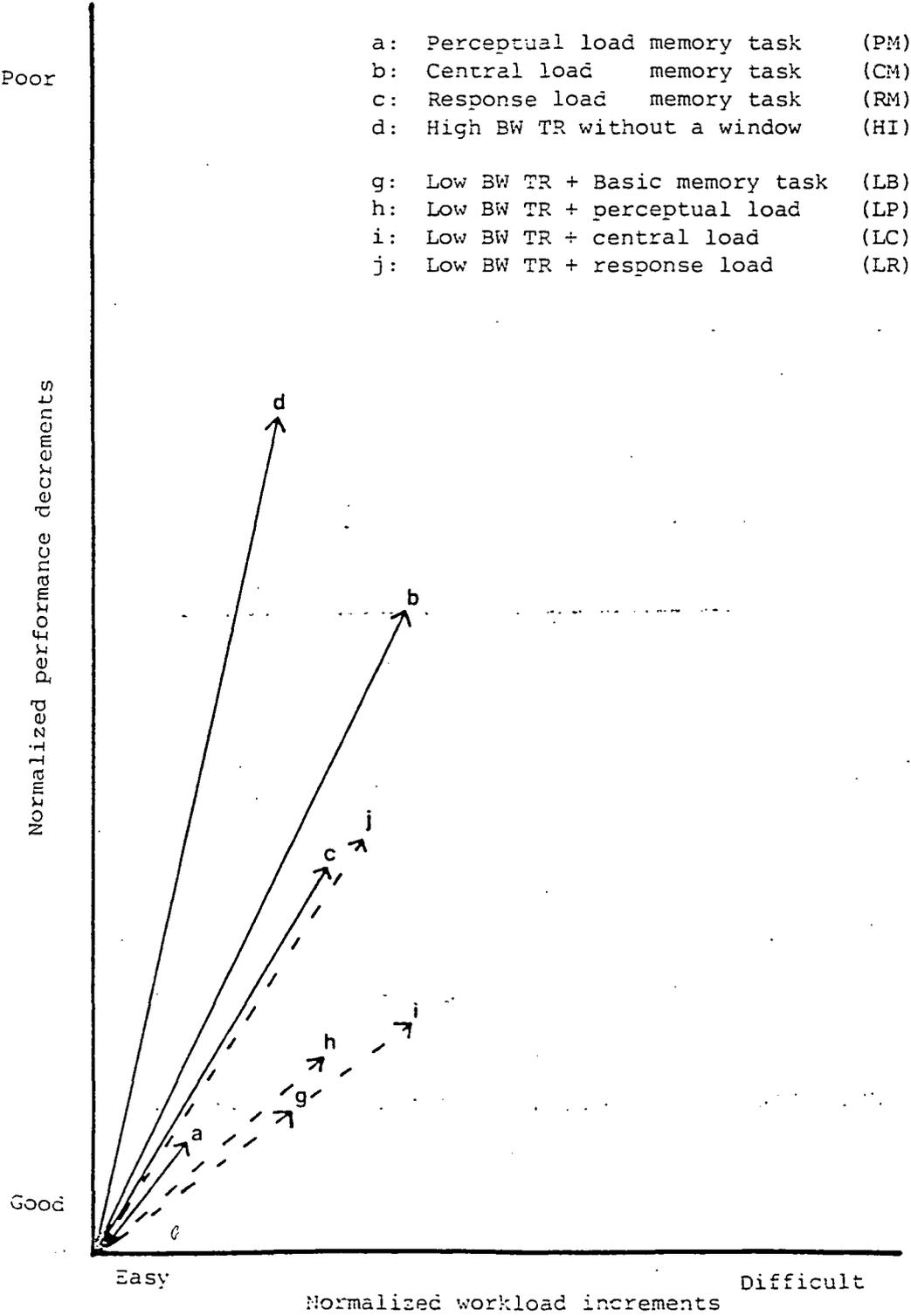


Figure 12. Performance and subjective workload of the conditions in Experiment 1

-----  
Insert Figure 13 about here  
-----

The dissociation in the data of Experiment 3 was shown when Vectors a and b in Figure 14 were compared. Vector a (DT) represents the effect of separating the display and Vector b depicts the effect of separating the control (CT) in a dual-axis tracking task. It was predicted that Vector b should point closer to the performance axis than Vector a. However, the result showed a contrary effect. Performance of the tracking task with separated control was much better than that of the same task with separated display. However, this advantage was underestimated by the difference in subjective workload ratings and the interaction was significant in the analysis of both standardized scores and decrement scores (DT vs. CT in Table 16).

-----  
Insert Figure 14 about here  
-----

In summary, only one paired comparison confirmed the prediction whereas the others showed a dissociation in contrast to the prediction. However, it should also be recalled that the additive factor logic failed to indicate any differential effect of the two tracking manipulations on perceptual and response load.

The number of display elements. Increasing the number of display elements is a manipulation that was predicted to increase the perceptual/cognitive load. This increase was predicted to be weighted heavily in the overall measure of subjective workload even if the added

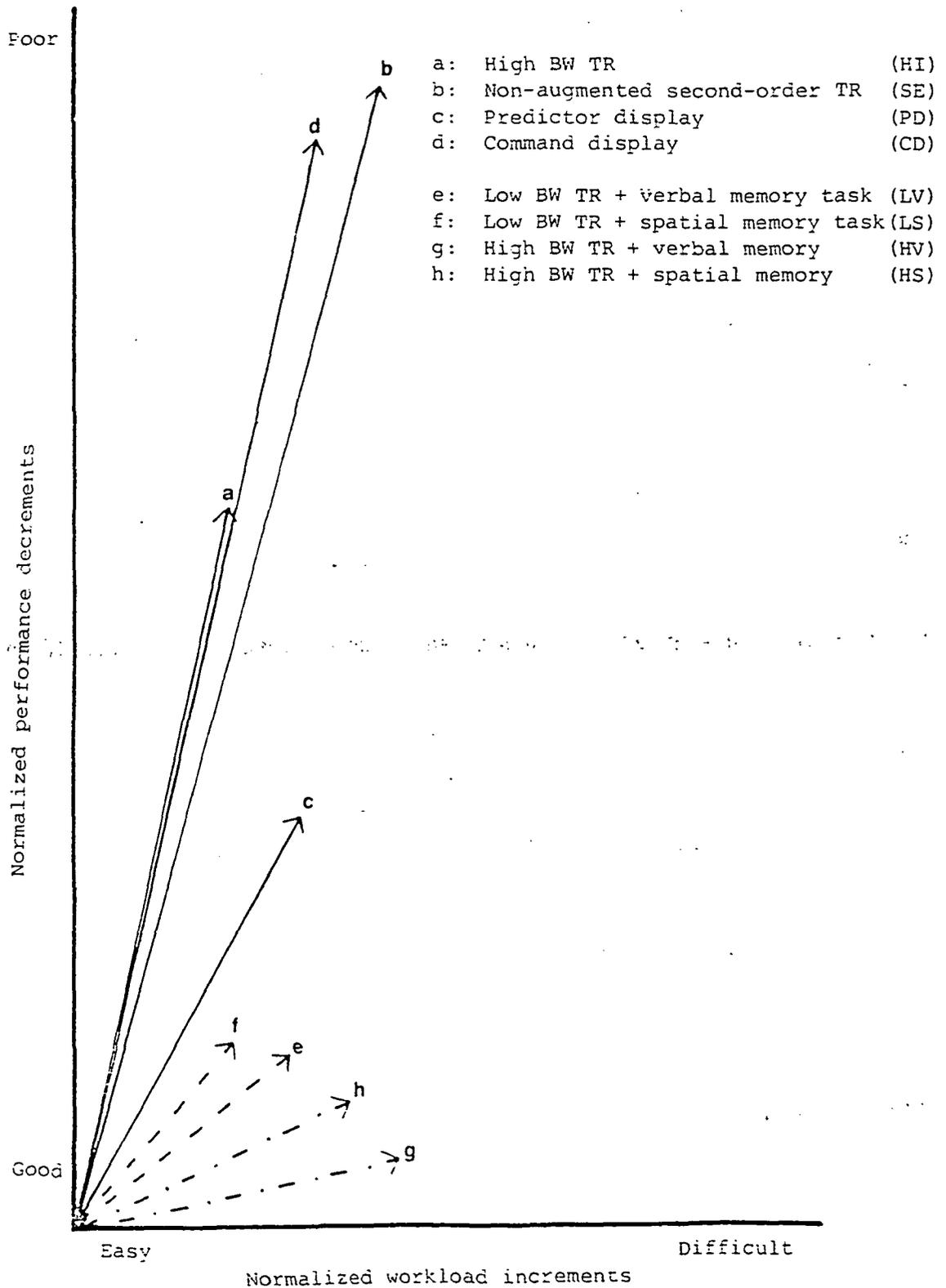


Figure 13. Performance and subjective workload of the conditions in Experiment 2

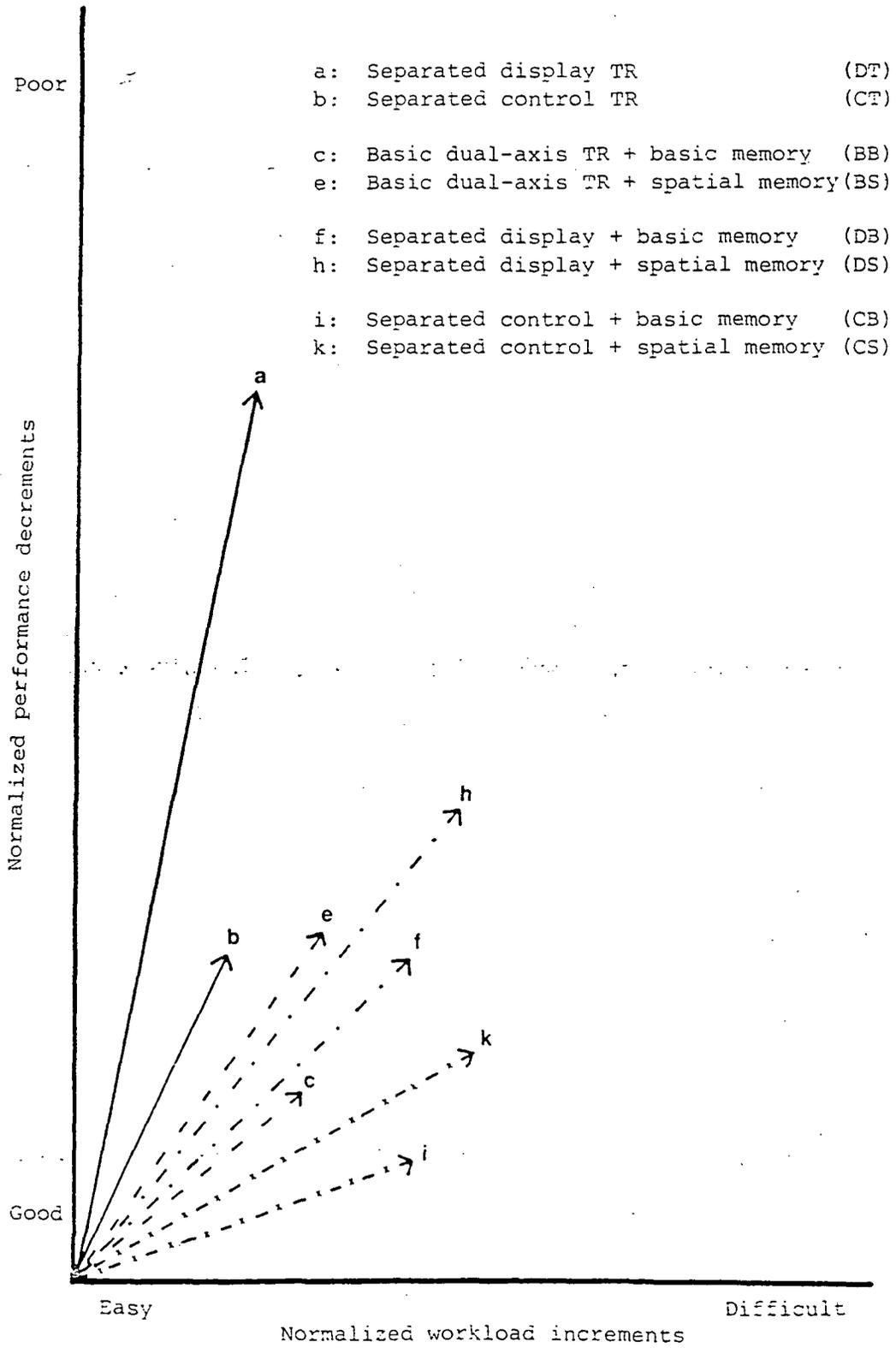


Figure 14. Performance and subjective workload of the conditions in Experiment 3

elements actually improved overall task performance. Therefore, the vector for the unaided second-order tracking task was anticipated to point relatively more toward the performance axis than the vector for the same task with a separate predictor. A similar result was also predicted when the vector for the object-like "command display" was compared with the vector for the two-element "predictor display".

Vectors for these three second-order tracking tasks are shown in Figure 13 (Vector b for the unaugmented tracking - SE, Vector c for tracking with a predictor display - PD, and Vector d for the command display - CD). In supporting the prediction, Vectors b and d (one element) point more toward the performance axis than Vector c (two elements). It appears that the predictor display improved tracking performance a great deal but subjective workload underestimated the amount of the benefit. The divergence between Vectors b and c (SE vs. PD in Table 16) was significant in the analysis of both standardized scores and decrement scores. The divergence between Vectors d and c (CD vs. PD in Table 16) was also supported by the statistics of all the three measures.

The command display was designed to integrate the additional control information into the cursor itself. It was presumed that this display would improve the tracking performance without increasing the subjective workload because there was no increase in the number of display objects. As shown in Figure 13, there was both a slight performance improvement and subjective workload reduction with the latter being greater than the former (compare vectors b and d). This difference produced a dissociation that was only significant in the analysis of effect sizes ( $F(1,14) = 5.33, p < .05$ ).

Single task difficulty vs. dual task competition. It was predicted that vectors for hard single task conditions would be closer to the performance axis than vectors for easy dual task conditions. The experimental comparisons that tested this prediction are listed in Table 17.

-----  
Insert Table 17 about here  
-----

This dissociation in the data of Experiment 1 is shown graphically in Figure 12. Vectors a (PM), b (CM), and c (RM) represent the three hard single memory tasks. Vector d (HI) shows the effect of increasing tracking bandwidth. Vector g illustrates the result of combining easy dual tasks (i.e., a low-bandwidth tracking with a basic memory search task - LB). As predicted, Vector g is further from the performance axis than any vector for the hard single tasks.

The results showed that the subjective workload of an easy dual task condition was higher than that of a hard single task even if performance was better in the former than in the latter condition. The statistics shown in Table 17 backed the findings. In comparing the dual task condition with the high-bandwidth tracking task (HI vs. LB in Table 17), all the three measures supported the dissociation between the workload measures. Two of the measures (standardized scores and decrement scores) supported the dissociation when the central load or the response load condition was compared with the easy dual task condition (CM vs. LB or RM vs. LB). One measure (standardized scores) supported the dissociation when the perceptual load condition was compared with the easy dual task condition (PM vs. LB).

The dissociation between the single task difficulty and the number of concurrent tasks in the data of Experiment 2 is shown in Figure 13. Vectors

Table 17 - Results of the interaction effect in ANOVA tests

Contrast: Single task difficulty vs. Dual task competition

<u>Comparison</u>	<u>Standardized Score</u>		<u>Effect Size</u>		<u>Decrement Score</u>	
	F value	P <	F value	P <	F value	P <
<u>Experiment 1</u>						
PM vs. LB	15.35	.002	1.26	-	2.62	.15
CM vs. LB	54.25	.00001	7.75	.01	41.96	.00001
RM vs. LB	18.40	.001	2.56	.20	23.52	.0004
HI vs. LB	35.85	.00001	16.27	.001	197.78	.00001
<u>Experiment 2</u>						
HI vs. LV	7.15	.02	5.41	.025	103.52	.00001
HI vs. LS	6.35	.025	4.74	.05	106.32	.00001
SE vs. LV	10.66	.006	10.52	.005	59.21	.00001
SE vs. LS	7.47	.02	9.62	.005	47.87	.00001
<u>Experiment 3</u>						
DT vs. BB	31.77	.0001	9.19	.01	86.26	.00001
CT vs. BB	2.26	.20	6.33	.03	5.86	.04

a and b show the effect of increasing bandwidth (HI) and control order (SE) respectively. Vectors e and f exemplify respectively the result of performing a low-bandwidth tracking task concurrently with a verbal memory task (LV) and with a spatial memory task (LS).

In supporting the prediction, Vectors e (LV) and f (LS) are closer to the subjective workload axis than Vectors a (HI) and b (SE). Thus, given an equal level of subjective workload, performance of a hard tracking task was much worse than that of an easy dual task condition. The statistics presented in Table 17 supported the findings. In comparing an easy dual task condition (Vector e or f) with a hard single task condition (Vector a or b), all the three measures supported the dissociation.

In Experiment 3 (Figure 14), the difficulty of a basic dual-axis tracking task (BT) was manipulated by separating the display elements or control (Vectors a and b respectively) or by adding an easy memory search task (Vector c - BB). Vectors a (DT) and b (CT) point to the left and above Vector c (BB), as predicted by the theory. It appears that given an equal performance level, subjective workload was higher in an easy dual task condition than in a hard tracking task. In comparing the dual task condition with the display-separated tracking task (DT vs. BB in Table 17), all the three measures supported the dissociation. Two of the measures (effect sizes and decrement scores) supported the dissociation when the control-separated tracking task was compared with the easy dual task condition (CT vs. BB in Table 17).

In summary, the dissociation between the single task difficulty and the number of concurrent tasks was found to be significant in most of the paired comparisons across the three experiments. The dissociation indicates that subjective workload of performing an easy dual task combination was much

higher than what the performance would indicate, relative to a hard single task.

Dual task competition for common vs. separate resources. It was predicted that performance would be adversely related to the amount of resource competition whereas subjective workload would be less sensitive to the difference in the amount of competition. That is, vectors for the high resource competition conditions should point more to the performance axis than vectors for the low competition conditions. The results of the ANOVA's testing of this hypothesis are presented in Table 18 and graphical presentations of the relevant data are shown in Figures 12 - 16.

-----  
 Insert Table 18 & Figure 15 about here  
 -----

The dissociations produced by the degree of resource competition in the data of Experiment 1 are shown in Figures 12 and 15. In Figure 12, Vector h represents the condition in which a low-bandwidth tracking task was performed concurrently with a perceptual load memory task (LP). Vector i depicts the corresponding condition with a central load memory task (LC), and Vector j with a response load memory task (LR). Vectors l, m, and n in Figure 15 represent the three corresponding conditions with a high-bandwidth tracking task (HP, HC, and HR).

The tracking task was inferred by the earlier additive factors analysis to demand response resources. Therefore, it was predicted that the vector for a dual task loading in the response load condition would point relatively more toward the performance axis than the vector for the loading in the perceptual/central condition (j vs. h & j vs. i in Figure 12, and n

Table 18 - Results of the interaction effect in ANOVA tests  
 Contrast: Dual task competition for common vs. separate resource

<u>Comparison</u>	<u>Standardized Score</u>		<u>Effect Size</u>		<u>Decrement Score</u>	
	F value	P <	F value	P <	F value	P <
<u>Experiment 1</u>						
LR vs. LP	17.8	.001	1.30	-	11.47	.005
LR vs. LC	2.63	.15	0.35	-	16.41	.002
HR vs. HP	22.07	.0005	0.92	-	17.54	.001
HR vs. HC	1.87	.20	0.79	-	17.12	.001
<u>Experiment 2</u>						
LS vs. LV	0.50	-	0.03	-	4.59	.05
HS vs. HV	0.13	-	0.09	-	6.10	.025
<u>Experiment 3</u>						
BS vs. BB	2.68	.20	0.90	-	12.22	.005
DS vs. DB	1.33	-	1.01	-	2.60	.15
CS vs. CB	0.29	-	1.15	-	0.90	-
BS vs. BC	12.31	.004	0.52	-	10.57	.007
DS vs. DC	14.82	.002	0.18	-	13.96	.003
CS vs. CC	7.25	.02	0.01	-	14.51	.002

ORIGINAL PAGE IS  
OF POOR QUALITY

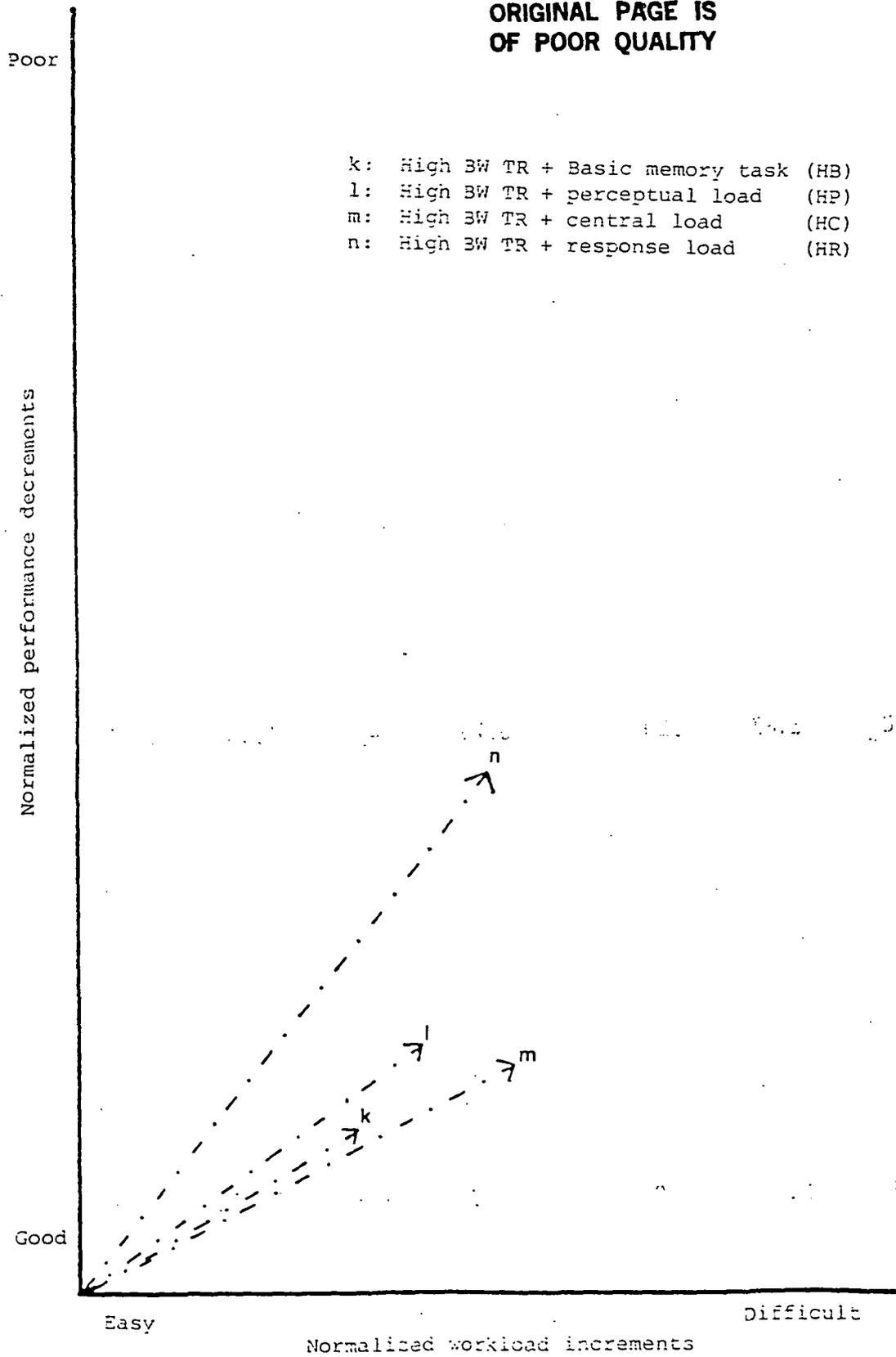


Figure 15. Performance and subjective workload of the conditions in Experiment 1 - High bandwidth tracking with a memory task

C-2

vs. l & n vs. m in Figure 15). As shown in these two figures, the prediction was confirmed.

The comparison for this dissociation could be made twice, once with the low-bandwidth tracking task and once with the high-bandwidth tracking task. The statistics shown in Table 18 supported the findings. In comparing the high competition condition with the condition in which the competition was low in perceptual load (LR vs. LP or HR vs. HP in Table 18), two of the measures (standardized scores and decrement scores) supported the dissociation. One measure (decrement scores) supported the dissociation when the high competition condition was compared with the condition in which the competition was low in central load (LR vs. LC or HR vs. HC in Table 18).

In Experiment 2, the tracking task was found to be time-shared more efficiently with a verbal memory task than with a spatial memory task, as predicted by the multiple resources model. Consistent with the proposed theory, Vectors e and g (a verbal memory task with a low-bandwidth tracking task and with a high-bandwidth tracking task respectively) in Figure 13 are closer to the subjective workload axis than Vectors f and h (a spatial memory task with a low-bandwidth tracking task and with a high-bandwidth tracking task respectively). Given the same performance level, subjective workload was higher when a tracking task was performed with a verbal memory task than with a spatial memory task. This dissociation was significant in the analysis of decrement scores (LS vs. LV and HS vs. HV in Table 18).

The dissociation produced by the difference in the degree of resource competition in the data of Experiment 3 is graphically shown in Figures 14 and 16. Vectors drawn in various types of dashed lines in Figure 14 represent the comparisons between different dual task conditions whose

memory set sizes were the same (tracking task with a basic verbal memory task vs. with a spatial memory task). Vectors in Figure 16 show the comparisons between the conditions whose memory set sizes were different (tracking task with a spatial memory task vs. with a central load memory task).

In both Figures 14 and 16, vectors drawn in simple dashed lines are the conditions in which the integrated dual-axis tracking task was performed concurrently with a memory task (Vector c for basic verbal memory task - BB, Vector d for central load memory task - BC, and Vector e for spatial memory task - BS). Vectors in a dot-dashed line (f, g, and h) represent the three corresponding conditions when the display-separated tracking task was time-shared with a memory task. Finally, vectors in an x-dashed line (i, j and k) depict the three corresponding conditions when the control-separated tracking task was executed concurrently with a memory task.

-----  
Insert Figure 16 about here  
-----

The results supported the prediction and replicated the results of Experiment 2 in which a single axis tracking task was employed. Vectors for the conditions in which a dual-axis tracking task was combined with a spatial memory task point relatively more toward the performance axis than vectors for the conditions in which the task was added to a basic verbal memory task (compare e vs. c, h vs. f, and k vs. i in Figure 14). Vectors for the spatial conditions also point relatively more toward the performance axis than vectors for the conditions in which the tracking task was added to a central load memory task (compare e vs. d, h vs. g, and k vs. j in Figure 16).

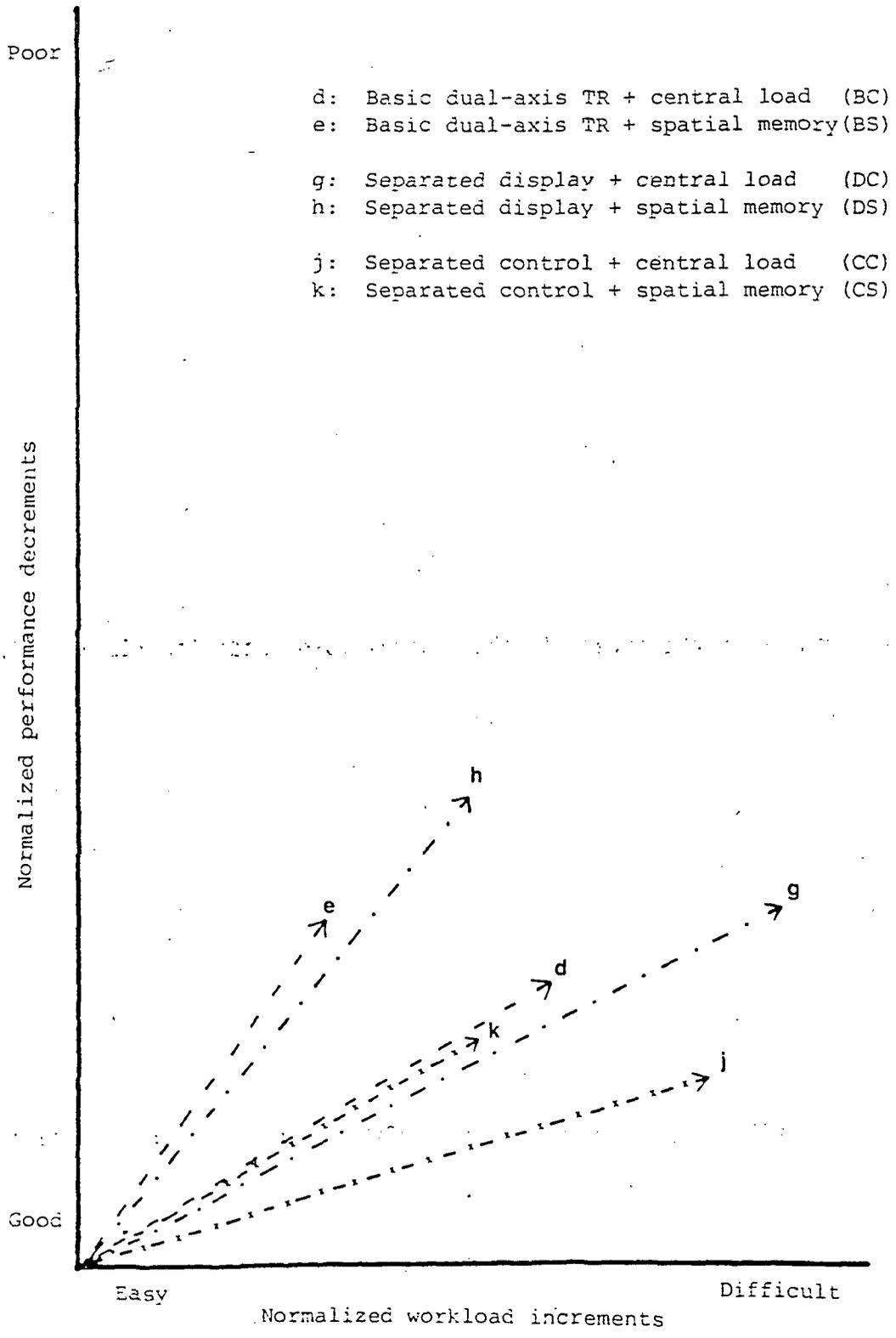


Figure 16. Performance and subjective workload of the conditions in Experiment 3 - Tracking with a spatial memory task vs. Tracking with a central load condition

It appears that given an equal performance level, subjective workload was lower when a dual-axis tracking task competed with a spatial memory task than when the tracking task was performed concurrently with a verbal memory task.

In Figure 14, the relative relationship between vectors, different in the amount of resource competition, was found in all three paired comparisons. However, only one comparison (e vs. c or BS vs. BB) showed a significant dissociation in the analysis of decrement scores ( $F(1,13) = 12.22, p < .005$ ). This dissociation was not found when the tracking task was in a format of separated display or separated control (f vs. h or j vs. i in Figure 14). All the three relevant paired comparisons in Figure 16 reached the significance level. In comparing the dual task effect in the spatial code with the effect in the central load condition (BS vs. BC, DS vs. DC, and CS vs. CC in Table 18), two of the measures (standardized scores and decrement scores) supported the dissociation.

In brief, the dissociation produced by the difference in the degree of resource competition between different dual task conditions was found in many paired comparisons across the three experiments. The results indicate that performance was effected more by the manipulation of resource competition than was subjective workload.

The effect of resource-investment factors. It was predicted that a motivation factor would improve performance at the cost of increasing subjective workload. This dissociation is shown by examining the three vectors drawn in solid lines in Figure 17 which represent the overall effects of the manipulations. In this figure, Vector d represents a high-bandwidth tracking task (HI), Vector e depicts the same task with a narrow

window (NA), and Vector f illustrates the tracking task with a wide window (WI). In the latter two conditions, subjects were motivated to invest more resources to reduce the presence of an unpleasant tone. In supporting the theory, Vectors e and f point more to the subjective workload axis than Vector d in Figure 17.

The results of the ANOVA's are listed in Table 19. In comparing the "narrow window" condition with the high-bandwidth tracking task, all the three measures supported the dissociation. Two of the measures (standardized scores and decrement scores) supported the dissociation when the "wide window" condition was compared with the condition without a window.

-----  
Insert Figure 17 & Table 19 about here  
-----

As shown in Figure 17, Vector e (NA) also points relatively more toward the subjective workload axis than Vector f (WI). The subjective workload of tracking with a narrow window was higher than with a wide window even though the difference in performance was small (performance actually degraded slightly with the narrow window). This divergence was significant in the analysis of standardized scores ( $F(1,14) = 8.81, p < .02$ ).

Examining individual subject's data, it was found that the width of the window had different effects on the performance of two different classes of subjects. For one group of subjects (which will be labelled Group 1), performance was worse in the "narrow window" condition than that in the "wide window" condition. For another group of subjects (Group 2), their performance was better in the "narrow window" condition than in the "wide window" condition.

ORIGINAL PAGE IS  
OF POOR QUALITY

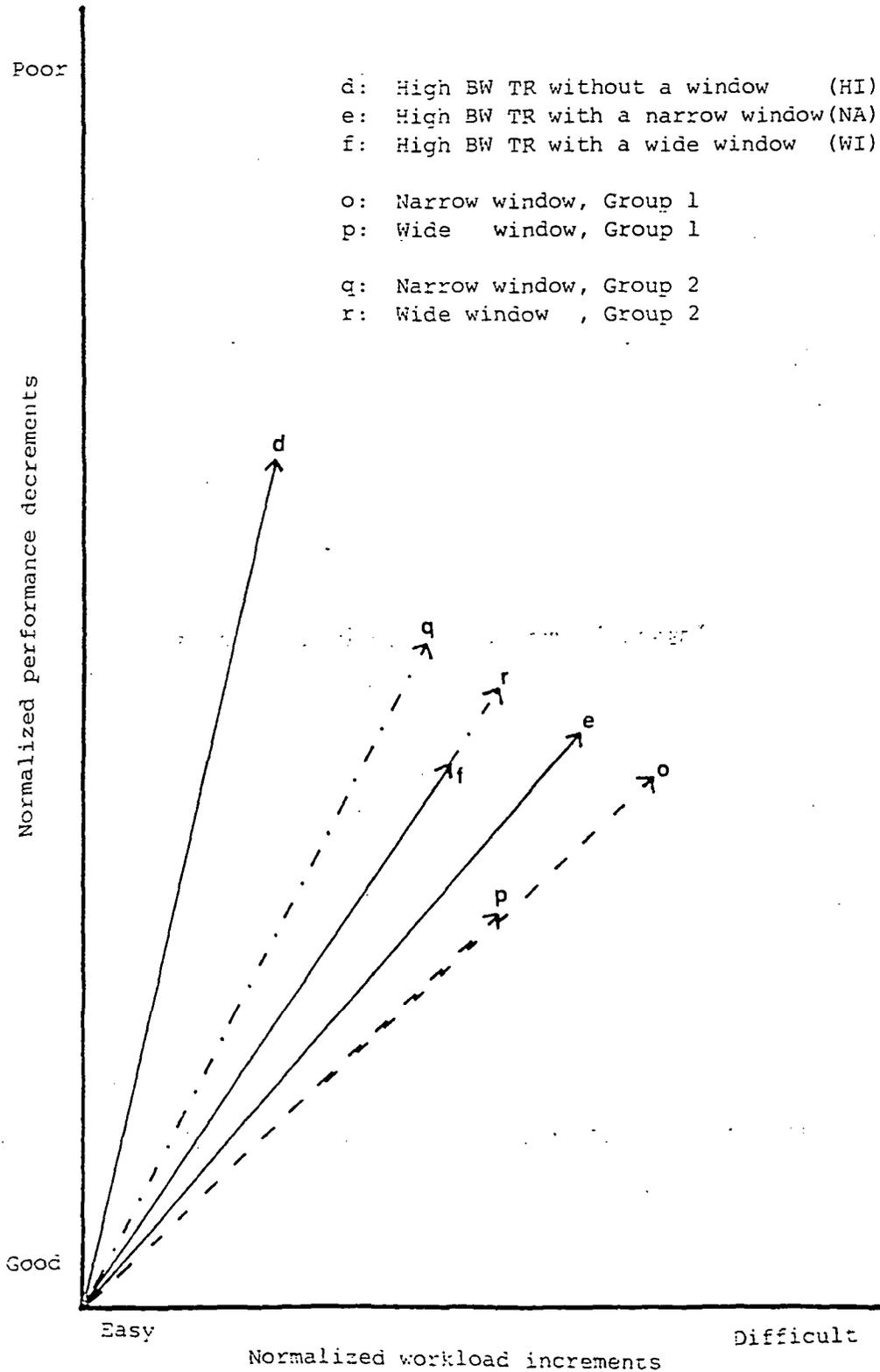


Figure 17. Performance and subjective workload of the high-bandwidth tracking conditions in Experiment 1

Table 19 - Results of the interaction effect in ANOVA tests

The effect of resource investment

<u>Comparison</u>	<u>Standardized Score</u>		<u>Effect Size</u>		<u>Decrement Score</u>	
	F value	P <	F value	P <	F value	P <
<u>Experiment 1</u>						
HI vs. NA	40.12	.00001	3.98	.10	33.84	.00001
HI vs. WI	40.31	.00001	0.12	-	38.54	.00001
NA vs. WI						
Overall	8.81	.02	0.18	-	2.88	.15
Group 1	1.33	-	0.39	-	0.05	-
Group 2	11.92	.015	0.96	-	9.18	.02
<u>Experiment 2</u>						
SE vs. PD	22.28	.0003	2.30	.20	65.17	.00001

Vectors o and p in Figure 17 represent the "narrow window" and the "wide window" conditions respectively for Group 1 subjects. The two vectors have the same slope and there was no dissociation between the two workload measures. For this group of subjects, they both performed more poorly and felt more loaded in the "narrow window" condition. Vectors q and r illustrate the "narrow" and "wide" window conditions respectively for the subjects in Group 2. Vector r points to the left and above Vector q. For this group of subjects, their performance was better in the "narrow window" condition, although they also felt more loaded. Two of the measures (standardized scores and decrement scores) supported this dissociation for Group 2 subjects (see Table 19). The nature of this difference in dissociation will be treated in the discussion section.

The dissociation between the unaided second-order tracking (SE) and the tracking task with a predictor display (PD) described previously in the context of Vectors b and c in Figure 13, may also be related to the effect of resource investment. When subjects performed a second-order tracking task with a predictor display, they were motivated to utilize the additional display information inherent in the predictor. Therefore, performance improved a great deal as the result of this display augmentation. However, because of the greater resource investment required to process the second display element, subjective workload underrated the benefit of the predictor display.

## Discussion

The primary objective of the present study was to identify sources of subjective ratings of difficulty and to determine how these ratings differ from performance. The multiple resources model (Wickens, 1984) was used as the theoretical framework for the investigation. Based upon this model, the processing characteristics of different tasks in a multidimensional space may be identified. Performance and subjective workload measures of these tasks may then be compared to examine how they relate to each other under various processing demands and why they dissociate.

Based upon the multiple resources model and previous findings (Wickens & Derrick, 1981) a theory of dissociation was proposed (Wickens & Yeh, 1983). According to the theory, dissociation occurs because (1) the demands on perceptual/cognitive resources, (2) the effect of resource-investment factors, and (3) the degree of resource competition, are all read unequally by the two types of measures. In the present study, three manipulations were employed to place demands on perceptual/cognitive resources. Two other manipulations were also employed, one to effect resource investment and the other to effect resource competition. Global workload ratings were used to represent the aggregate demands from all the components and were then compared with performance to test the hypotheses drawn from the proposed theory.

The effect of these manipulations on the structure of perceived difficulty was also disclosed by a multidimensional scaling approach. The result of this analysis was used to provide converging evidence with previous findings (Derrick & Wickens, 1984; Yeh & Wickens, 1984) which deny

that the two measures dissociate simply because subjective workload itself is unidimensional, as Gopher and Braune (1983) implied. If the processing characteristics effect the load experienced by the subjects, the demands on perceptual/cognitive resources should be unveiled as one of the subjective dimensions. The regression analysis of the prediction of overall workload from specific rating scales, was used to present another confirmation that perceptual/mental effort was an important component of subjective workload. The information about the components of subjective workload provides guidance to the workload practitioners concerning what rating scales should be chosen to assess workload.

When subjects perform a task or dual task combination, both performance and subjective ratings may be used to evaluate the load of the task imposed upon the human information processing system. Performance, measured by reaction time, accuracy, or tracking error, directly reflects the internal nature of the processing characteristics. On the other hand, perceived workload, estimated by some rating scales, represents the introspection of the load experienced by the processing system. The processing characteristics are the common ground that underlies both types of measures imposed by the task. Processing characteristics, agreed to be multidimensional, can best be identified by a measure which preserves the vector properties in the multidimensional space. Performance, a vector quantity, has been recognized as such a measure (see Wickens, 1984 for a review). Subjective ratings can also unveil the multidimensionality via the multidimensional scaling approach.

Why would performance and subjective workload dissociate if they both estimate the load of the same processing characteristics? Could it be that

the structure of subjective workload really differs from the structure of processing resources that underlies the performance? According to Gopher and Braune (1983), subjective workload follows the pattern of the most restricted undifferentiated pool while performance follows the pattern of a multiple resources model. Using the multidimensional scaling approach, Derrick and Wickens (1981, 1984) have demonstrated that the structure of perceived task difficulty is related to the structure of processing resources as portrayed by the multiple resources model. The results of the present study also supported this corresponding relationship. It appears that subjective workload is indeed multidimensional although its multidimensional characteristics are not a direct analog of the multidimensional characteristics of performance.

Given that the structure of subjective workload corresponds to the structure of processing resources, the multiple resources model provides a powerful theoretical basis for studying the sources that produce dissociation between the two measures. Use of the multiple resources model permitted a manipulation of task difficulty by placing different demands and competition on certain resource dimensions (e.g., input and output modality, codes of central processing, or processing stage). In the present study, the demands of a single task and competition between two tasks were systematically manipulated employing two types of tasks, tracking and a Sternberg memory search task. These manipulations were carried out via different means: (1) excessive loads were placed on a set of resources (e.g., second-order tracking), (2) the demand on a specific processing stage was varied (perceptual/central and response load memory task), and (3) the degree of competition for a processing stage or a processing code was manipulated in dual task conditions.

The additive factors method was used to confirm the processing demands of various tracking tasks. This information was then used to verify the degree of resource competition in different dual task conditions. The results of the additive factors analysis indicated that (1) A low-bandwidth single axis tracking task was time-shared efficiently with a verbal memory task, but not with a spatial memory task or a response load memory task. Thus, the demands of a tracking task were conclusively verified to be both spatial and response-related. (2) The manipulation of increasing bandwidth produced an additive effect (i.e., it did not enlarge the effect of adding a low-bandwidth tracking task to a memory task). This result replicates the finding of Wickens and Derrick's study (1981) that increasing bandwidth demands resources that are not specific to any processing stage or processing code. (3) The predictor display demanded more perceptual resources in comparison with an unaugmented display. (4) Separating the display cursor or the control input in a dual-axis tracking task did not impose extra loads on any specific processing stage or code.

### MDS results

With the processing demands of the single tasks diagnosed, the processing characteristics or the interference pattern of a dual task condition may be understood in terms of the demand and competition on each resource dimension. This information provided a basis for examining the relationship between performance and subjective workload. Prior to the investigation of the relationship, a measure was needed to represent the subjective introspection of workload. The subjective measure of overall workload was found to correlate with three of the four dimensions in the data of Experiment 1 (Table 13), with the two dimensions revealed from the

MDS data of Experiment 2 (Table 14), and with all the three dimensions disclosed from the data of Experiment 3 (Table 15). Therefore, this scale was used to indicate the aggregate weighting on the components in the multidimensional space and to represent the subjective introspection of the experienced load. Overall workload ratings of various processing characteristics were then compared with performance to identify the sources of dissociation.

According to the proposed theory of dissociation (Wickens & Yeh, 1983), dissociation occurs because certain demands are read differently by the two measures. Performance reflects every aspect of the processing dynamics whereas subjective workload is postulated to be sensitive to the amount of aggregate resource investment and be dominated by the demands for perceptual/cognitive resources. The salience of the perceptual/cognitive demand in subjective introspection was partially supported by the structure of subjective workload revealed from the data. All the three subjective dimensions disclosed from the MDS data of Experiment 3 were related to the demands on perceptual/cognitive resources. Two subjective dimensions (the processing stage and the demands for perceptual/cognitive resources) revealed from the data of Experiment 1 were also associated with this processing characteristic. Moreover, the high regression weights of the perceptual/mental ratings on the prediction of overall workload also suggest that the demand for perceptual/cognitive resources is a salient component of overall workload.

Thus, it is clear that the subjective experience of mental workload is strongly influenced by the perceptual/cognitive demands. The issue of whether these demands also represent a source of dissociation with performance will be discussed below.

### Discussion of dissociations

Three types of manipulations were employed in the present study to place demands on perceptual/cognitive resources. (1) The number of concurrent tasks: The difficulty of a single task was increased by adding another easy task. In such a condition, many subsystems in the multidimensional space are utilized and executive management is necessary to coordinate the time-sharing between the two tasks. This cost of concurrence was assumed to impose an extra load on perceptual/cognitive resources. (2) The number of display elements: This manipulation was imposed by separating or integrating the additional predictive control information with the cursor in a second-order tracking task in Experiment 2. The manipulation on the number of display elements was also employed by separating the display cursor of a dual-axis tracking task in Experiment 3. When the number of elements was greater, the display required more perceptual/cognitive resources. (3) Perceptual/cognitive difficulty: The difficulty of a basic Sternberg memory search task was increased in Experiments 1 and 3 by placing an extra load on the perceptual or central stage of processing in order to elevate the demands on these resources.

Some dissociation effects between performance and subjective workload were found as the consequence of the different demands on perceptual/cognitive resources imposed by different manipulations.

(1) Single task difficulty vs. dual task competition. According to the proposed theory, both manipulations will elevate subjective workload and damage performance. However, performing two concurrent tasks will place an additional load on perceptual/cognitive resources because the executive management is necessary to process and coordinate two tasks. Hence,

subjective workload will overestimate the performance decrements in easy dual task conditions, relative to difficult single task conditions.

Replicating previous findings (Derrick & Wickens, 1984; Wickens & Yeh, 1983), a strong dissociation was found. As summarized in Table 17, this dissociation was shown in various comparisons across different task configurations and across the three types of methods from which comparable measures were derived. Performance decrements were lower when doing a hard single task than when doing an easy dual task combination, but subjective workload was higher in the latter condition.

(2) The number of display elements. According to the theory, increasing the number of display elements will impose more demands on perceptual/cognitive resources and hence elevate the subjective workload. The effect of the predictor display was one example of this dissociation. When a second-order tracking task was performed with a predictor, two elements (cursor and predictor) were processed and coordinated. Therefore, demands on perceptual/cognitive resources were higher and subjective workload underrated the benefit of the predictor on the tracking performance.

The weak dissociation produced by the command display was also in accord with the effect of the number of display elements. Performance improved slightly, but not significantly with the command display which did not include a separate display element. However, the control information was provided by the integrated object-like symbol which reduced the requirement of generating lead. Thus, the demands for perceptual/cognitive resources decreased and subjective workload was lower in the command display condition than what its performance would indicate.

One dissociation was found to be contrary to the prediction. In Experiment 3, two display elements were processed in the display-separated dual-axis tracking task. On the other hand, only an integrated cursor was processed in the control-separated tracking condition. The results from Experiment 3 indicated that subjective workload was similar between these two conditions even though performance was worse in the display-separated condition. This contrary dissociation may result from the lack of specificity of resource demands of the display-separated condition. The results of the additive factors analysis showed that unlike the predictor display, the separate tracking cursors placed no more demands on perceptual/cognitive resources than did an integrated display. Therefore, whether the number of display elements will produce a dissociation may depend upon whether the manipulation imposes extra demands on perceptual/cognitive resources. Further research is needed to systematically investigate the effect of integral/separable perceptual features on the relationship between performance and subjective workload measures.

(3) Perceptual/cognitive difficulty. Imposing demands on the perceptual/central resources and on the response-related resources of a task should equally effect performance, but the former was predicted to affect subjective workload more than the latter manipulation. Unlike the previous two manipulations, this one failed to provide conclusive evidence for the dissociation.

The only evidence for this dissociation was observed when the perceptual load condition was compared with the response load memory task in Experiment 1. A reverse dissociation was shown when the central load condition and response load memory task were compared (i.e., subjective

workload underrated the performance decrements in the central load memory search task). This reverse effect may result from the high cognitive load in the central load memory task in which three sets of alpha-numeric strings were held in working memory. Under a high cognitive load condition (Eggemeier et. al., 1982), subjective workload is a less sensitive measure than performance.

In Experiment 2, subjects underestimated the performance decrements of the second-order tracking task, in comparison with the high-bandwidth tracking task condition. This dissociation is in opposition to the previous finding. In interpreting the previous result, Wickens and Yeh (1983) suggested that subjective workload was higher for the second-order tracking task than what its performance would indicate because this task required subjects to predict the acceleration and hence required more perceptual/cognitive resources. The conflicting results between the two studies may be due to a difference in procedure across the two studies. In the previous study, the system gain was the same in both the low and high order conditions, but in the present study the system gain of the second-order tracking task was much lower than that of the high-bandwidth tracking condition. The magnitude of control inputs must be very large in order to compensate for the low gain in the second-order tracking task and hence require subjects to allocate more response resources to the task. Future research is necessary to resolve this conflicting evidence.

The MDS data of Experiment 1 showed that the resource dimension of processing stages was one of the subjective dimensions of task difficulty (Dimension 1 in Figure 8). This information, combined with the results of manipulating the demands on stage-related resources, suggests that the demands on different processing stages may be weighted equally in subjective

introspections of workload. This finding contradicts the proposed theory in which two basic assumptions were made: (1) subjective introspection reflects information heeded in working memory, and (2) working memory is represented primarily by the perceptual/central resources. Based upon these two assumptions, the theory predicted that the demands on response resources would not be accurately read by subjective ratings.

Results from dual task studies suggest that the validity of the second assumption remains uncertain. Wickens (1978) reviewed many dual task studies and pointed out that "short-term memory processes draw upon a general resource pool that is available to and used by all concurrent tasks, whatever their modality and processing stage...." (p248). Therefore, the demands on the response resources, in addition to the demands on perceptual/central resources; may be heeded in working memory and hence reflected in subjective introspection of mental workload. Dissociation between the two workload measures may occur only when the aggregate resource demands in working memory are different between the two tasks.

In addition to placing demands on perceptual/cognitive resources, two manipulations were executed to investigate the other two sources of dissociation proposed by the theory. (1) Resource-investment factors: Intrinsic task-related features were employed to induce more resource investment through an imposition of a predictor element on a second-order tracking task in Experiment 2 and an imposition of an error window on a high-bandwidth tracking task in Experiment 1. (2) Dual task competition for common vs. separate resources: Competition for a processing stage or processing code was manipulated to produce different degrees of resource competition in various dual task conditions.

Dissociation was found as the result of both manipulations.

(1) Resource-investment factors. Factors which induce more resource investment, in order to improve performance, were predicted to increase subjective workload. On the other hand, factors that reduce the amount of invested resources will decrease subjective workload but deteriorating performance. In the present study, two types of factors were used to induce greater investment of resources. The first one was the predictor display in a second-order tracking task. By offering more precise information, this predictor display would lead the subjects to use that information and hence invest more resources in the processes. On the other hand, superior performance would occur as the benefit of the more precise information. This effect in fact was obtained. The dissociation was confirmed when the "predictor display" condition was compared with the unaugmented display.

Motivational variables are another type of factor that may induce greater investment of resources. Using pay bonus as an incentive, Vidulich and Wickens (1983) demonstrated that an external motivational factor may produce this sort of dissociation. Under the bonus condition, subjects' performance was better than that under the no bonus condition. But, subjects felt equally or more loaded when doing the task in the bonus condition. Tulga (1978) showed that a subjective variable could also produce dissociation. When the load exceeded the information processing ability, subjects could not reach the objective demand. When the discrepancy between the actual performance and the objective criterion was large, subjects tended to lower their performance criterion. The change of performance criterion produced worse performance but lowered the perceived workload.

In the present study, another type of motivational variable was employed via intrinsic task-related features. A window (independent of the width of the acceptable error range) in a high-bandwidth tracking task induced the subjects to invest more resources to reduce the presence and duration of an unpleasant tone. Greater investment of resources under a "window" condition resulted in superior performance in comparison with the "no-window" condition and dissociation between the two workload measures occurred.

The difference between Group 1 and Group 2 subjects indicated that there were individual differences in fulfilling the objective demand. In general, Group 2 subjects performed better than Group 1 subjects in the 2nd-order tracking tasks. Group 2 subjects (better trackers) invested more resources to match their performance with the objective demand in the "narrow" window condition. Their performance was better than the tracking performance in the "wide" window condition, but they felt more loaded.

Group 1 subjects (worse trackers), in contrast, gave up trying to achieve the objective demand and lowered the performance criterion in the "narrow" window condition. Like Group 1 subjects, they felt more loaded in the "narrow" than in the "wide" window condition, but unlike Group 1, their performance was worse in the narrow window condition. The association between the workload measures for this group of subjects was a phenomenon different from what was observed by Tulga (1978). In Tulga's study, performance deteriorated when subjects gave up achieving the objective demand but the subjective workload decreased.

All the subjects felt more loaded under the narrow window condition than under the wide window condition for two reasons: (1) The degree of required precision influences subjective workload (Moray, 1982). In the

"narrow" window condition, the unpleasant tone was presented more often and longer than in the "wide window" condition. Subjects were aware that they were required to invest more resources to meet the objective demand and hence they reported a higher workload value. (2) The high frustration level experienced by failing to meet the performance demand in the narrow window condition.

The dissociation produced by resource-investment factors is actually the propelling force behind the use of subjective ratings. Many researchers argue that performance itself does not always convey valid information of mental workload (Johannsen et. al., 1979; Sheridan, 1981; Moray, 1982). Subjects may simply put in more effort to prevent performance decrements resulting from an increase in task load. Therefore, performance may not show any change with the elevating load. The results of the present study and the previous ones (Vidulich & Wickens, 1983; Tulga, 1978) suggest that the dissociation between performance and subjective workload is inevitable in some conditions. When more resources are induced by an incentive (bonus) or intrinsic task-related features (predictor and the window), performance improves at the cost of an increasing subjective workload. When subjects lose their motivation and reduce the amount of invested resources, their performance drops but they feel less loaded (Tulga, 1978). When the required precision level produces a high frustration level, subjects may feel more loaded even though they lower the performance criterion (Group 1 subjects in the narrow window condition).

(2) Dual task competition for common vs. separate resources. From dual task studies (see Wickens 1981 for a review), it has been shown that performance is adversely effected by the amount of resource competition between tasks. However, independent of the amount of resource competition,

the total demand for all resources or the function of the executive management may be constant. Therefore, subjective workload, being more sensitive to the aggregate resource investment, was predicted to underestimate the difference in performance decrements between dual task conditions with competitive or separate resource demands.

Results of previous studies (Derrick & Wickens, 1984; Wickens & Yeh, 1983) showed that a strong dissociation occurs when there is a salient difference in the amount of resource competition between two dual task conditions. When a tracking task was time-shared with an auditory task (demands are spread well over many resources), performance was better than when the task competed with itself (maximum competition). However, subjective workload was not sensitive to such a difference in resource competition.

In the present study, both a tracking task and a memory search task were performed in all the dual task conditions. The competition was restricted to the processing stage (Experiment 1) and the code of central processing (Experiments 2 and 3). Therefore, the difference in the amount of resource competition was less salient than the manipulations in the previous studies. Nevertheless, dissociation was still found in many task comparisons. Independent of the locus of competition (processing stage or code), performance decrements in the high resource competition conditions, in comparison with the low competition conditions, were greater but underrated by subjective workload estimates.

#### Strength of dissociations

Among the five aspects of dissociations, some are stronger than the others and are found in many paired comparisons across the three experiments

and across the three methods from which comparable units are derived. The ratio (AD/PD) of the number of significant dissociations actually observed to the number of potential dissociations (number of comparison \* 3 types of methods), was quite different among the five different dissociation phenomena. The strongest dissociation occurs between the single task difficulty and the dual task combination, whose AD/PD is 26/30. The second strongest one is the dissociation produced by the resource-investment factors (AD/PD is 7/9). The dissociation produced by the degree of resource competition is found at 15/36. In all three cases, it is important to note that none of the paired comparisons refute the prediction (i.e., showed a significant dissociation in the opposite direction).

The AD/PD ratio produced by the number of display elements is 6/12. One comparison (dual-axis tracking with display separated vs. tracking with control separated) showed a contrary dissociation which may result from the finding that separate cursors in this task did not place additional demands on perceptual/central resources. The resource dimension of processing stages appears not to be a source of dissociation. Two out of 12 potential dissociations confirm the prediction but 5 out 12 refute it.

In summary, subjective workload, sensitive to the amount of aggregate resource investment, dissociated from performance when the investment was increased to improve performance. Dissociation also occurred when two dual task conditions had different degrees of resource competition. Furthermore, the demand on perceptual/cognitive resources was a salient component of the subjective perception of mental workload. When the demand on these resources was raised by imposing a time-sharing requirement or by increasing the number of display elements, a dissociation occurred. When the demands were imposed on different processing stages, the demands on response

resources and on perceptual/central resources were read equally by the performance and subjective workload.

### Subjective workload scales

Due to the multidimensionality of the internal processing, it is plausible that certain vector properties of subjective ratings are not preserved by a global scale. In the MDS data of Experiment 1, overall workload ratings correlated with three subjective dimensions whereas response effort ratings were tied to all the four dimensions. This result suggests that the response effort ratings had indeed picked up some information that was not registered in the overall workload ratings.

If the response effort ratings did indeed respond to task characteristics that were not reflected in the global workload ratings, it is possible that the former might also show less dissociation from performance. To test whether the dissociation would diminish by using the response effort ratings, several comparisons were retested and the data are presented in Table 20. Among these comparisons, the weak dissociation in Experiment 1 produced by loading different processing stages (perceptual load vs. response load or vector a vs. c in figure 12) lost its significance. It seems, not surprisingly that the demands of the response load memory task were tapped by the response effort ratings better than by the overall workload ratings.

In contrast, reflecting the overall workload, the response effort ratings were not sensitive to the competition between two tasks for response resources. The dissociation produced by the degree of competition for stage-related resources (low-bandwidth tracking with a perceptual load memory task vs. the tracking task with a response load memory task) remained

Table 20 - Results of the interaction effect from ANOVA's on  
response effort ratings

Dissociation phenomenon: Perceptual features affect subjective workload  
more than it influences performance

<u>Comparison</u>	<u>Standardized Score</u>		<u>Effect Size</u>		<u>Decrement Score</u>	
	F value	P <	F value	P <	F value	P <
PM vs. RM	.61	-	1.07	-	2.13	-
SE vs. PD	24.28	.0002	1.31	-	54.45	.00001
DT vs. CT	17.49	.002	1.73	-	46.27	.00001

Dissociation phenomenon: Single task difficulty vs. dual task competition

<u>Comparison</u>	<u>Standardized Score</u>		<u>Effect Size</u>		<u>Decrement Score</u>	
	F value	P <	F value	P <	F value	P <
RM vs. LB	26.26	.0002	0.49	-	13.62	.003
HI vs. LB	37.22	.00001	16.79	.001	190.47	.00001
HI vs. LV	2.74	.12	11.44	.001	61.15	.00001
SE vs. LV	5.88	.03	17.52	.001	48.50	.00001
DT vs. BB	18.68	.001	4.38	.05	66.08	.00001

Table 20 (continued) - Results of the interaction effect from ANOVA's on response effort ratings

Dissociation phenomenon: Dual task competition for common vs. separate resources

<u>Comparison</u>	<u>Standardized Score</u>		<u>Effect Size</u>		<u>Decrement Score</u>	
	F value	P <	F value	P <	F value	P <
LR vs. LP	12.53	.005	1.18	-	16.52	.002
HR vs. HP	16.29	.002	1.03	-	22.99	.0003
HR vs. HC	0.05	-	0.61	-	14.79	.002

Dissociation phenomenon: The effect of a motivational factor

HI vs. NA	27.69	.0001	3.15	.10	25.89	.0005
-----------	-------	-------	------	-----	-------	-------

significant. Furthermore, the dissociation produced by the number of concurrent tasks or by a motivational variable was also significant. Hence, although response effort ratings do reflect differences in the subjective workload, these differences are not sufficient to account for the major sources of performance-subjective workload dissociation.

If it is the case that information on some subjective dimension is not fully preserved by the global ratings, what scales should be used by workload practitioners?

The most popular multiscale procedure of measuring subjective workload is the Subjective Workload Assessment Technique (Reid, et. al., 1981). In this procedure, rankings are made on three scales: time pressure, stress level, and effort load. Averaged ranking data are fitted by an additive combination rule via a conjoint measurement methodology. This rule then determines how rankings on the three scales should be combined and the combined value is used to indicate subjective workload. However, the choice of these scales is atheoretical and certain problems exist in the procedure and the logic of this assessment technique (Boyd, 1982).

In theory, resource dimensions could be used to assess subjective workload since the structure of information processing is the common basis for both performance and subjective workload. However, the exact structure of subjective workload depends on the range of processing patterns in the multidimensional space expanded by the set of tasks employed in a study.

Processing characteristics are more likely to be revealed in the structure of subjective dimensions when they are very distinguishable among many tasks in a set. For example, four tasks were employed in Derrick and Wickens' study (1981, 1984): a critical tracking task, a visual search task, an auditory Sternberg task, and a tone judgment task. In a set of 18

tasks (four easy, four hard, and 10 dual task conditions), about half of the tasks required resources from the visual modality and the other half demanded auditory resources. Consequently, input modality was one of the subjective dimensions. The amount of resource competition was also quite diverse in different dual task conditions. When the tracking task was paired with itself, the competition was the maximum. When the tracking task was combined with an auditory Sternberg task or a tone judgment task, the competition as defined by the multiple resources model was close to zero. The demands were well spread over input modality (visual vs. auditory), processing codes (verbal vs. spatial), and response mode (manual vs. speech). As a result, resource cost and competition was revealed as another subjective dimension of workload.

In the context of other sets of tasks (Yeh & Wickens, 1984), the code of central processing and the aggregate resource cost (which depended upon the number of subsystems in the multidimensional space engaged in the processing, the amount of demands on each subsystem, the locus of demands on each resource dimension, and the distribution of the competition among the reservoirs) were found to be the subjective dimensions of task difficulty.

In the present study, the amount of demand for perceptual/cognitive resources was a major manipulation across different tasks. The outcome, as expected was that this processing characteristic was disclosed from the MDS data of Experiments 1 and 3. Converging evidence from the previous studies (Derrick & Wickens, 1984; Yeh & Wickens, 1984) and from the present study indicates that the demand of a single task and competition between dual tasks on resource dimensions are the cornerstones of the structure of subjective workload. Whether a particular resource dimension will be revealed from the data depends upon the processing patterns imposed by the

set of tasks employed in a study. Since the exact structure of subjective workload in a set of tasks can only be disclosed post hoc, how may a workload practitioner choose the scales a priori? Which scales should be selected to assess the load of different processing characteristics?

Apparently, a global scale may be used to tap the aggregate demands and variances of all the components, independent of the structure. In Derrick and Wickens' (1984) study, the perceived effort was correlated with all three disclosed subjective dimensions. In the present study, overall workload was a scale that correlated with most subjective dimensions. The task complexity scale and the modified Cooper/Harper scale also correlated with all of the subjective dimensions revealed from the data of Experiments 2 and 3. Ratings on these latter two scales had significant correlations with the overall workload ratings (Table 11). It seems that any global scale may be chosen to tap the aggregate demands. The overall workload or perceived effort scale is recommended for the high face validity.

Should specific ratings be collected if the overall workload ratings encompass the load on all of the components? Since the exact structure of subjective workload depends on the set of tasks, it is always possible that specific demands are not fully represented by global ratings. The results from the first two experiments illustrate this point. The response effort ratings did not contribute significantly in accounting for the variance of the overall workload ratings (Table 12). However, response effort ratings correlated with the two dimensions disclosed from the data of Experiment 2. Response effort ratings also correlated with all of the four subjective dimensions from the data of Experiment 1 whereas the overall workload ratings correlated with only three dimensions. When the response effort ratings were used in the analysis, the weak dissociation produced by

comparing the perceptual and response load memory tasks vanished. It seems that one dimension was better represented by the response effort ratings than by the overall workload ratings. Therefore, specific ratings may be useful to tap some demands left out by global ratings.

The choice of specific scales should be based upon the nature of the important components of subjective workload. The correlations between subjective dimensions and unidimensional ratings (Tables 13-15) as well as the results of the regression analysis (Table 12) suggest four scales: perceptual/mental effort, response effort, stress, and time-pressure. The first three scales were also found to be salient components of task difficulty or overall workload in Vidulich and Wickens' study (1983). Time-pressure ratings were not recommended by Vidulich and Wickens (1983) because these ratings were insensitive to the effect of some manipulations. Therefore, perceptual/mental effort, response effort, and stress are recommended to assess the demands on specific aspects.

Given that the demand for working memory was an important component of subjective workload, immediate ratings may be subject to some phenomenal aspects that are task-specific and less stable. Results from a previous study (Yeh & Wickens, 1984) and from the present study indicated that post trial ratings were in general less reliable than post session ratings even though both were fairly reliable.

The difference between rating reliabilities may occur because the effect of specific task features is more likely to diminish within the context of all of the just-performed tasks. For example, an alpha-numeric string is particularly easy to remember because it has a special meaning for a subject. When this subject is asked to rate the load immediately after he performs the task, his ratings will be very different from the ratings of

the same condition with another memory set. On the other hand, ratings collected within the context of all of the just-performed tasks are more relative. Subjects have a chance to retrospect the experiences of doing each task. They may compare the differences in general processing characteristics and rank order the ratings for each condition. It is not argued that immediate ratings are hence less valid. It is only suggested that post session ratings, in the long run, may provide relatively stable and reliable ratings of the processing characteristics. Immediate ratings are more vulnerable to specific task features that fluctuate from time to time. If a workload practitioner is concerned with the relative workload from operating different systems, ratings within the task context are recommended.

### Conclusion

In conclusion, the results from the present study support the view that the multiple resources model provides a common theoretical basis for understanding both the objective and the subjective aspects of mental workload. Through this theoretical model, processing characteristics can be diagnosed in terms of the demands on resource dimensions. This knowledge presents a powerful framework for understanding the relationship between workload measures and for determining why workload measures dissociate.

The dissociation occurs when demands are read unequally by the two workload measures. Performance, a vector measure, is influenced by every aspect of processing. On the other hand, subjective workload is more sensitive to the amount of aggregate resource investment. Furthermore, the demand for perceptual/cognitive resources was shown to be a salient

component from both the MDS analysis of the hidden structure and the regression analysis of the underlying components of overall workload.

The executive management engaged in processing and coordinating two tasks was found to be the most potent factor that drove subjective workload. A hard single task was favored by subjects even though their performance was better in a dual task condition. A workload practitioner who relies upon subjective ratings rather than performance, may be biased to choose a non-optimal system that requires operators to perform just one task rather than the system that demands dual task performance.

Factors that induce different amounts of resource investment were also a source of a strong dissociation. The effect of these resource-investment factors exposes one indispensable dilemma existing in workload measurement. On the one hand, any factor that discourages resource investment may decrease subjective workload at the cost of deteriorating system performance. On the other hand, encouraging better performance may increase subjective workload. Furthermore, intrinsic task-related features that aid performance will also elevate subjective workload. A workload practitioner who relies upon subjective ratings, may be led to abandon predictor displays that produce better performance than conventional displays.

The amount of resource competition may also produce a dissociation. Dual task performance is adversely effected by the amount of competition between two tasks, but subjective workload may not show the difference, or may indicate that the difference is trivial. A workload practitioner should be aware of this potential dissociation when a system is selected by subjective workload. Finally, the resource dimension of processing stages, in contrast, was found not to be a source of dissociation. This resource dimension was disclosed as one of the subjective dimensions. Moreover, the

weak and reverse dissociations suggest that demands for response-related resources and perceptual/central demands may be equally weighted in subjective introspection of workload. It is suggested that the demands on response resources may also be heeded in working memory and hence be reflected in subjective workload.

The relationship between workload measures has been a puzzle in workload assessment. Given the same manipulations, workload measures may correlate in some conditions and dissociate in other conditions. What is needed in this field is a theoretical model to understand the perplexing relationships (Moray, 1982) and how they diversify as the processing characteristics alter. Derrick's study (Derrick & Wickens, 1981, 1984) is the first attempt to study the relationship between workload measures from the framework of the multiple resources model. The studies by Wickens and Yeh (1983) and Vidulich and Wickens (1983) have confirmed the utility of such an approach. The present study is a continuing effort to refine our knowledge of this relationship. Although some aspects (e.g., stress) were not examined, many aspects of processing characteristics were manipulated to identify potential sources of dissociation. There is still a need for research to complete our knowledge and to determine the appropriate methods for collecting subjective ratings. It is believed that the results from the present study provide a foundation for future research.

## References

- Ackerman, P. L. (in press). Multivariate evaluation of workload, resources, and performance in dual task data. Proceedings of the 28th Annual Meeting of the Human Factors Society.
- BMDP-81. (1981). Berkeley, CA: University of California Press.
- Boyd, S. P. (1982). The use of conjoint analysis for interval subjective scaling of mental workload. Proceedings of the 8th Psychology in the Department of Defense Symposium, 334-339.
- Derrick, W. (1981). The relationship between processing resource and subjective dimensions of operator workload. Proceedings of the 25th Annual Meeting of the Human Factors Society, 532-536.
- Derrick, W. & Wickens, C. D. (1984). A multiple processing resource explanation of the subjective dimensions of operator workload (Tech. Rep. EPL-84-2/ONR-84-1). Urbana-Champaign: University of Illinois, Engineering Psychology Laboratory.
- Daryanian, B. (1980). Subjective scaling of mental workload multi-task environment. Unpublished master's thesis, MIT, Cambridge, MA.
- Ericsson, K., & Simon, H. A. (1980). Verbal reports as data. Psychological Review, 87, 215-250.

- Eggemeier, F. T., Crabtree, M. S., & Reid, G. B. (1982). Subjective workload assessment in a memory update task. Proceedings of the 26th Annual Meeting of the Human Factors Society, 643-647.
- Garner, W. R. (1974). The processing of information and structure. Potomac, MD: Erlbaum.
- Gartner, W. B., & Murphy, M. R. (1976). Pilot workload and fatigue: A critical survey of concepts and assessment techniques (NASA TND-8356). Moffett Field, CA: NASA Ames Research Center.
- Gopher, D., & Braune, R. (1983). On the psychophysics of workload: Why bother with subjective measures? Proceedings of the 2nd Symposium on Aviation Psychology, Ohio State University, Columbus, OH.
- Hart, S. G., Childress, M. E., & Hauser, J. R. (1982). Individual definitions of the term "workload". Proceedings of the 8th Psychology in the Department of Defense Symposium, 334-339.
- Hauser, J. R., Childress, M. E., & Hart, S. G. (1982, June). Rating consistency and component salience in subjective workload estimation. Paper presented at the 18th Annual Conference on Manual Control, Dayton, OH.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. Journal of Educational Statistics, 6, 107-128.

- Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments. Psychological Bulletin, 92, 490-499.
- 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.
- Johannsen, G., Moray, N., Pew, R., Rasmussen, J., Sanders, A., & Wickens, C. D. (1979). Final report of experimental psychology group. In N. Moray (Ed.), Mental workload: Its theory and measurement. New York: Plenum Press, 101-114.
- Kahneman, D., & Triesman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman, J. Beatty, & R. Davies (Eds.), Varieties of attention. New York: Academic Press, 29-62.
- Kramer, A. F. (1983). The processing of stimulus attributes: Evidence for dual-task integrality. Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign.
- Kruskal, J. B., & Wish, M. (1978). Multidimensional scaling series: Quantitative applications in the social sciences. Beverly Hills, CA: Sage.
- Moray, N. (Ed.) (1979). Mental workload: Its theory and measurement. New York: Plenum Press.

- Moray, N. (1982). Subjective mental workload. Human Factors, 24, 25-40.
- Micalizzi, J., & Wickens, C. D. (1980). The application of additive factors methodology to workload assessment in a dynamic system monitoring task (Tech. Rep. EPL-80-2/ONR-2). Urbana-Champaign: University of Illinois, Engineering Psychology Laboratory.
- Murphy, M. R., McGee, L. A., Palmer, E. A., Paulk, C. H., & Wempe, T. E. (1978). Simulator evaluation of three situation and guidance displays for V/STOL aircraft zero-zero landing approaches. Proceedings, 1978 IEEE International Conference on Cybernetics and Society, 563-571.
- Navon, D., & Gopher, D. (1979). On the economy of the human processing system. Psychological Review, 86, 214-255.
- Norman, D., & Bobrow, D. G. (1975). On data-limited and resource-limited processing. Journal of Cognitive Psychology, 7, 44-60.
- Reid, G. B., Shingledecker, C. A., & Eggemeier, F. T. (1981). Application of conjoint measurement to workload scale development. Proceedings of the 25th Annual Meeting of the Human Factors Society, 522-526.
- Sanders, A. F. (1979). Some remarks on mental workload. In N. Moray (Ed.), Mental workload: Its theory and measurement. New York: Plenum Press, 41-77.

Sadry, D. L., & Wickens, C. D. (1982). The effect of stimulus-central processing-response compatibility and resource competition on pilot performance (Tech. Rep. EPL-82-1/ONR-82-1). Urbana-Champaign: University of Illinois, Engineering Psychology Laboratory.

Sheridan, T. B. (1980). Mental workload - What is it? Why bother with it? Human Factors Society Bulletin, 23 (2), 1-2.

Sheridan, T. B., & Simpson, R. W. (1979). Toward the definition and measurement of the mental workload of the transport pilots (Tech. Rep. DOT-OS-70055). Cambridge, MA: MIT Man-Machines Systems Laboratory.

Sternberg, S. (1966). High-speed scanning in human memory. Science, 153, 652-654.

Tulga, M. K. (1978). Dynamics decision making in multi-task supervisory control: Comparison of an optimal algorithm to human behavior. Unpublished doctoral thesis. MIT, Cambridge, MA.

Tulga, M. K., & Sheridan, T. B. (1980). Dynamics decision and workload in multitask supervisory control. IEEE Transactions on Systems, Man, and Cybernetics, SMC-10, 217-231.

Wickens, C. D. (1973). The structure of attentional resources. In R. Nickerson & R. Pew (Eds.), Attention and performance VIII. Englewood Cliffs, NJ: Erlbaum, 239-258.

- Wickens, C. D. (1981). Processing resources in attention, dual task performance, and workload assessment (Tech. Rep. EPL-81-3/ONR-81-3). Urbana-Champaign: University of Illinois, Engineering Psychology Laboratory.
- Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman, J. Beatty, & R. Davies (Eds.), Varieties of attention. New York: Academic Press, 63-102.
- Wickens, C. D., & Derrick, W. (1980). The processing demands of higher order manual control: Application of additive factors methodology (Tech. Rep. EPL-80-2/ONR-2). Urbana-Champaign: University of Illinois, Engineering Psychology Laboratory.
- Wickens, C. D., & Derrick, W. (1981). Workload measurement and multiple resources. Proceedings, 1981 IEEE International Conference on Cybernetics and Society, 600-603.
- Wickens, C. D., Mountford, S. J., & Schreiner, W. S. (1981). Time-sharing efficiency: Evidence for multiple resources, task-hemispheric integrality, and against a general ability. Human Factors, 23, 211-229.
- 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, 244-248.

Wierwille, W., & Williges, R. (1978). Survey and analysis of operator workload assessment techniques (Tech. Rep. S-78-101). Blacksburg, VA: Systemetric, Inc.

Wierwille, 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.

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

Yeh, Y-Y., & Wickens, C. D. (1984). An investigation of the dissociation between subjective measures of mental workload and performance (Tech. Rep. EPL-84-1/NASA-84-1). Urbana-Champaign: University of Illinois, Engineering Psychology Laboratory.