N86-32980

Levels of information processing in a Fitts law task (LIPFitts)

Kathleen L. Mosier
University of California
Berkeley, CA

Sandra G. Hart
NASA-Ames Research Center
Moffett Field, CA

ABSTRACT

State-of-the-art flight technology has restructured the task of human operators, decreasing the need for physical and sensory resources, and increasing the quantity of cognitive effort required, changing it qualitatively. Recent technological advances have the most potential for impacting the contempory pilot in two areas: performance and mental workload. In an environment in which timing is critical, additional cognitive processing can cause performance decrements, and increase a pilot's perception of the mental workload involved. The effects of stimulus processing demands on motor response performance and subjective mental workload are examined in the current study, using different combinations of response selection and target acquisition tasks. The information processing demands of the response selection were varied (e.g., Sternberg memory set tasks, math equations, pattern matching), as was the difficulty of the response execution. Response latency as well as subjective workload ratings varied in accordance with the cognitive complexity of the task. Movement times varied according to the difficulty of the reponse execution task. Implications in terms of real-world flight situations are discussed.

INTRODUCTION

Typical aircraft control tasks require, in some proportion, three types of resources: physical, sensory, and cognitive processing. The job of the contemporary pilot seldom demands strenuous physical effort, other than staying awake and alert on long or fatiguing flights. It requires a small degree of sensory effort, such as reading gauges and listening to warning clackers, etc., and a continually increasing amount of cognitive processing (e.g., calculations, instrument comparisons, decisions) that often must be performed quickly with little margin for error. Flying tasks that were once accomplished by sensory means now demand more sophisticated mental effort, since displays present integrated and refined information rather than raw data. In addition, the quality of cognitive effort required has been redefined. For example, digital readouts are replacing analog gauges, requiring number processing on the part of the operator rather than a quick glance to ascertain that the arrows on several dials are pointing in the same expected direction. Even the task of finding an airport has evolved to a cognitive processing task because of the need to use localizers, instrument approaches, etc. in addition to looking out of the window.

Finally, when timing is critical, extra cognitive processing may increase the time to respond to a signal (Hart, Sellars, & Guthart, 1984), causing a performance decrement. Even a task as simple as moving left or right in response to a command is more difficult and time-consuming when the information is presented linguistically (e.g., "RIGHT") rather than spatially. For example, Hart et al. (1984) found differences in reaction time (RT) performance based simply on a directional arrow (>) versus a linguistic command (R/L). They also found an additional 40-msec lag in RT when subjects were required to process the size and distance of a target in addition to the directional cue.

Sternberg (1975) and others found RT performance differences depending on the number of items a subject was required to remember and search through (the memory set) before responding as to whether another stimulus (the probe) was or was not a member of the memory set. It is reasonable to expect that these response decrements found in controlled, laboratory experiments that involve relatively minor levels of cognitive processing would, if anything, be exascerbated in a more realistic flight situation, with the potential for life-threatening situations.

Accompanying the demand for a thinking, vigilant, analytical pilot has been a concern over the amount of cognitive load that is placed upon the operator as well as the type of load. Since most of resources currently being tapped are cognitive, it is quite likely that an increase in the complexity of the cognitive demands of a task would have a measurable effect on the pilot's perception of the workload involved. Physical workload is relatively easy to predict and measure, although one is limited by observable behaviors, such as the movement of arms, hands, fingers, and legs, and eyes. Overload results in physical fatigue, injury, or inability to perform a task. Mental workload (i.e., how much a pilot can be expected to process, remember, or analyze in a given time span) is, however, much more elusive. Although mental workload is becoming more and more precisely defined, individual interpretations of the concept itself, as well as its various components, have hindered accurate measurement.

The model for the tasks used in the present study was the "FITTSBERG" paradigm (Hartzell, Gopher, Hart, Dunbar, & Lee, 1983), which combines, serially, a FITTS target acquisition task (Fitts and Peterson, 1964) with a SternBERG memory task (Sternberg, 1975). The decision of which two targets to acquire is based on the results of a Sternberg-type memory search. A series of experiments has been conducted employing variations of this paradigm to investigate the relationship between stimulus processing demands and motor response performance (e.g., Hart et al., 1984). In the original study (Hartzell et al., 1983), subjects were given a choice of two targets, one to the right and one to the left of center. The difficulties of the target acquisitions were indexed (ID) according to Fitts' law (Fitts and Peterson, 1964). The direction of the movement was based on whether or not the probe stimulus was (right) or was not (left) a member of the Sternberg memory set. Memory sets of 1, 2, or 4 letters were used. When compared with performance on a single target task, RT for the combined "Fittsberg" task was sensitive to the additional cognitive processing requirements of the Sternberg memory tasks. As expected, the impact of response selection complexity did not extend into the movement phase (from initiation of response to target capture criterion). Movement times (MT) were not significantly different than for target acquisitions without a response selection requirement.

In subsequent studies, the workload of the two component tasks (target acquisition and response selection) together was judged to be considerably less than the summed workload of each task done separately. The subtle differences in RT for directional versus linguistic cues continued to be reliable; as was the 40-msec increase in response selection time (RST) with the addition of a target acquisition (TA) task. In a recent study (Staveland, Hart, & Yeh, 1985), it was found that different measures of performance (e.g., RT, RST, MT) selectively reflected different portions of the Fittsberg task, and could be manipulated independently. The workload ratings reflected the average workload within a block of trials (exhibiting no primacy/recency effects of trial difficulty) and integrated the workloads imposed by both selection and execution components.

The present experiment expanded the Fittsberg paradigm to include many other types of information processing, including pattern and rhyme recognition, time estimation, and mathematical problem solving. It also varied the types of information in the memory sets (eg. categories, numerical values, and words, as well as individual letters) and the memory interval (immediate, delayed). The difficulty of the cognitive task that determined movement direction ranged from simple, single-step decisions (e.g., whether or not two simultaneously appearing letters were identical) to relatively complex decisions that required several steps (e.g., solving a complex arithmetical equation and comparing the result to the numerical value of the memory set function).

Current research has focused on the subjective experience of mental workload, either by itself or in combination with performance and physiological measures (Wierwille and Casali, 1983) as the most valuable estimate of load. Multi-dimensional approaches to subjective workload measurement take into account the idea that the experience of workload is a cumulative effect of three (e.g., stress, mental effort, and time pressure) or more factors (Reid, Shingledecker, Nygren, & Eggemeier, 1981), and that the same elements objectively occurring in the same proportions may lead to different estimations of workload from different performers. To account for individual interpretations of factors associated with workload, a system has been devised to combine ratings for each factor with weights reflecting the subjective importance given to that factor (Hart, Battiste, & Lester, 1984). This weighting system, used in conjunction with nine different elements of workload and an overall workload evaluation, was used in the present study.

The goal of the present study was to relate performance and workload changes associated with 10 different information processing tasks. In terms of performance:

1) The difficulty of a response selection task is reflected in its latency (RST), decision reversals and percent correct. Initiation of a target acquisition is measured by RT.
2) The difficulty of a target acquisition is reflected in MT, but not in the initial RT (single alternative) or RST (two alternatives).
3) If the effect of response selection difficulty extends into the movement phase, MTs will increase.
4) If information processing for response selection and initiating

response execution are performed serially in the Fittsberg (FB) condition: RST(FB) = RST + RT.
5) If processing is accomplished in parallel: RST(FB) = RST or RT, whichever is greater (implying that no extra time is required for the processing of the additional task).
6) If response selection and initiation of target acquisition overlap, but each requires some unique processing: RST + RT > RST(FB) > RST or RT.

With respect to the subjective ratings of workload (WL):
1) If subjective workload is affected by task complexity, workload ratings will parallel RST and RT differences.
2) If FB imposes more workload than simple response selection tasks, WL(FB) > WL(RS). In this case, either a) workload ratings for the combined tasks will equal the sum of the component task workload ratings [WL(FB) = WL(RS) + WL(TA)]; or b) because of a certain amount of functional overlap, the workload of the combied tasks will be equal to the load imposed by the response selection task plus some non-overlapping part of TA [WL(FB) = [WL(RS) + WL(TA)] * C, where C < 1.0 and C > 0.5].
3) If no additional workload is imposed by the TA task, then FB workload will be equal to the rating of RS or of TA, whichever is greater [WL(FB) = WL(RS) or WL(TA)].

It was hypothesized that: a) RSTs would mirror task complexity; b) information processing for RS and TA would progress essentially concurrently; c) control reversals and percent correct would be affected by response selection task complexity only; d) MTs would reflect target ID only; e) subjective workload ratings would also coincide with task complexity; and f) the extra demands of the TA condition would result in slightly higher, but not additive, workload ratings.

METHOD

Subjects

Nine subjects, ranging in age from 18 to 40, served as paid participants. All of them had been previously trained on different versions of the Fittsberg task that were not used in this experiment (i.e., Sternberg memory sets of one, two, and four with a Fitts target acquisition).

Apparatus

The experimental chamber contained a chair 85cm from a 23-cm monochrome monitor. On the right or left arm of the chair (depending on the handedness of the individual) was a two-axis joystick used for making RT, RST and TA responses. Workload-related ratings were obtained with a slide pot and enter-button on the non-dominant arm rest. An additional switch was mounted on the non-dominant arm rest for response selection in right-target-only and left-target-only conditions. An Apple II computer was used for target generation and data collection (10-msec resolution).

Experimental tasks

Ten response selection tasks, involving several levels of cognitive effort, were presented alone and in combination with a Fitts TA task. The pattern match (PM) task was selected as the basic response selection task for the TA control condition, due to its relatively simple processing and memory demands. For most tasks, an answer that was "yes" or "greater" prompted a movement to the right (and acquisition of the target on the right on TA trials). Tasks required no memory, recent (previous trial) memory, or "long term" memory. Each was performed first as a simple response selection task, then as a FB task in combination with a TA. Table 1 illustrates the experimental tasks.

Reaction time (RT) and RST were defined as 2% deflection of the joystick. Three IDs were used for TA, computed in accordance with Fitts' law. Width varied from 5 to 20 pixels, and target distance from 60 to 128 pixels [ID(2.52) = 40/60; ID(4.19) = 7/64 or 14/128; ID(5.67) = 5/128]. Except for the control conditions, the three target IDs were randomly presented within each block of 24 trials. Movement time (MT) was calculated from stick deflection to a steadiness criterion (keeping the cursor in the target).

Feedback

In all tasks (except time estimation) descriptive feedback about correctness and RST was given after each trial, and, where applicable, MT. The time criteria for each feedback phrase remained constant throughout tasks and conditions. Norms for intervals used in providing feedback were derived from earlier studies. Descriptive adjectives comparing current performance to the norms ranged from "truly dismal" to "fantastic".

Subjective rating scales

Nine elements of workload were rated: task difficulty, time pressure, own performance, physical effort, mental effort, frustration, stress, fatigue, and activity type (skill- or knowledge-based). Before beginning the experiment, subjects were asked to evaluate the importance of each element to overall workload, compared to every other element, by making 35 pairwise comparisons. The final weight of each factor ranged from 0 (never considered more important than another factor) to 8 (considered more important than any other factor) (Hart et al., 1984). At the end of each experimental block, subjects were asked to rate their experience on each of the nine workload factors, as well as to give an overall workload rating, on 10 bipolar rating scales.

Procedure

After completing the factor weightings, subjects were given an introduction describing the study and the tasks they would be performing, accompanied by demonstration trials. They were given two practice and one experimental block for each task, followed by ratings, in a previously-determined, counterbalanced order. All subjects performed the tasks in the response-selection-only mode first. Prior to performing the TA condition, they were given two practice and one experimental block of trials for each of the control conditions: PM + easy (ID = 2.52) TA (PME); PM + hard (ID = 5.67) TA (PMH); and PM + easy/med/hard TA, right TA only (PMR) or left TA only (PML). A block consisted of 24 trials.

Table 1

Experimental Tasks

EXPERIMENTAL TASK	MEMORY SET	PROBE	MOVEMENT
PATTERN MATCH (PM)	—	T W	MATCH - RIGHT NO MATCH - LEFT
ODD/EVEN (O/E)	—	9	EVEN - RIGHT ODD - LEFT
RHYME (RYM)	—	PAIN SANE	RHYME - RIGHT NO RHYME - LEFT
TIME ESTIMATION (T)	—	5 SEC	WAIT 10 SEC - RIGHT OR WAIT 5 SEC - LEFT
WITTENBORN (W)	—	3 6 1	SUM OF OUTER #S > MIDDLE # - RIGHT LESS - LEFT

EXPERIMENTAL TASK	MEMORY SET	PROBE	MOVEMENT
RECENT MEMORY CHANGING MS			
GREATER/LESS (G/L)	6	8	GREATER THAN PREVIOUS # - RIGHT LESS - LEFT
NUMBER SET (SET)	5 4 0	3 6 2	ANY #S IN PREVIOUS SET - RIGHT NONE - LEFT
DELAYED MEMORY			
EQUATION (EQ)	(3+3) ÷ 2	(6x2) ÷ 6	GREATER THAN MS - RIGHT LESS - LEFT
STERNBERG RHYME (SRYM)	TRIAL	VILE	RHYME WITH MS - RIGHT NO RHYME - LEFT
STERNBERG TIME ESTIMATION (TS)	10 - 2	12 - (2x3)	GREATER THAN MS - WAIT 10 SEC - RIGHT LESS - WAIT 5 SEC - LEFT

RESULTS

The data collected for each task were 1) RT, RST, or time duration prior to deflection (for time estimation tasks); 2) MT (where applicable); 3) percent correct; 4) control reversals (e.g., second thoughts about response selection, and 5) bi-polar workload ratings. Several analyses of variance were performed across experimental conditions for each measure: percent correct; RT for TA-only tasks; RST for response-selection-only tasks; and RST, control reversals, and MT for FB tasks. Time estimation tasks were analyzed separately, since RSTs were equal to the duration of 5- or 10-sec time productions. Most of the tasks were also grouped and analyzed by type: 1) control condition (PM, PME, PMH, PMR, PML); 2) math functions (G/L, W, EQ); 3) time estimation (T, TS); and 4) rhyme (RYM, SRYM).

In general, RST was shown to be very sensitive to response selection difficulty, $F(7, 56) = 22.33$, $p<.01$. The addition of the target acquisition task further enhanced this effect (Figure 1). Weighted workload ratings exhibited this sensitivity as well, $F(23, 184) = 8.75$, $p<.01$ (Figure 2). Right/left response differences were not significant, except for tasks in which direction of movement was determined by a yes/no choice. In this case, "no" responses were somewhat slower. Movement time, as expected, was not affected by response selection difficulty or the number of alternative targets. A significant effect was found across all tasks for percent correct, $F(23, 184) = 10.46$, $p<.01$.

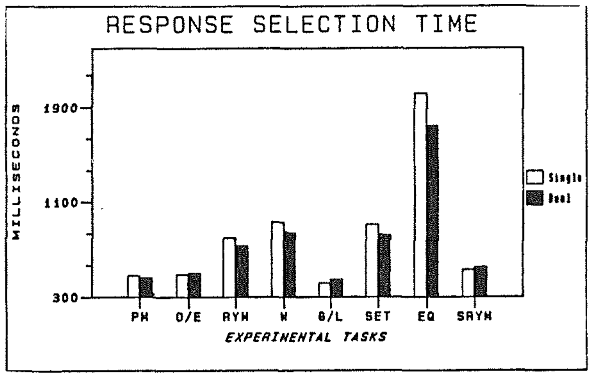

Figure 1. Response selection times for all tasks.

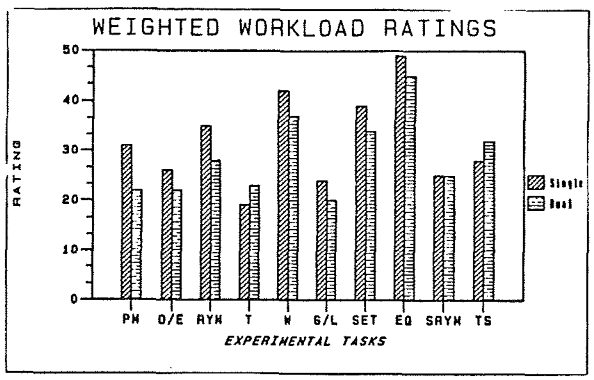

Figure 2. Weighted workload ratings for all tasks.

<u>Control Conditions</u>

Within the pattern match conditions, the effects of several variations of the TA portion of the FB task were examined, i.e., keeping the target ID constant (PME, PMH); keeping the direction of movement constant (PML, PMR); and removing the response selection requirement from target acquisition (PML, PMR). Results of the pattern match condition followed the expected pattern. No significant differences were found for RT or percent correct, and direction of movement (right versus left) did not have a significant effect.

Movement time differences were found as a function of target ID, as predicted by Fitts' law (Fitts and Peterson, 1964): average MT for easy targets (PME) was .695 msec; for hard targets (PMH), 1.065 msec (Figure 3). A significant interaction was found between PME/PMH and right/left, $F(1,8) = 9.18$, $p < .05$; i.e., the easy/hard MT differences were somewhat more pronounced for right targets than for left targets.

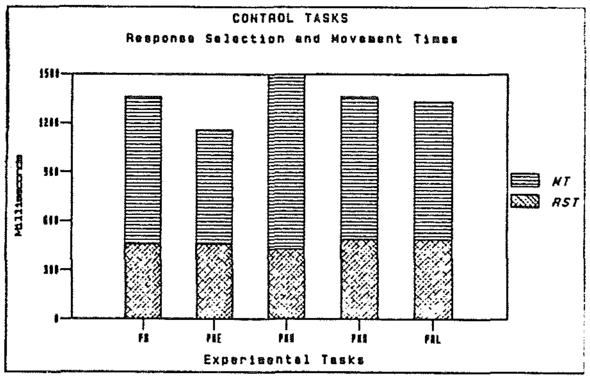

Figure 3. Response selection and movement times for control condition.

In PMR and PML conditions, two RT measures were taken: one for the RS task (a button press); and one for the RT following target appearance (joystick deflection). Responses to the target alone, involving no cognitive processing task, were predictably faster than for any of the cognitive tasks, and were not affected by target difficulty. When one element, either target side (R/L) or ID (E/H), was held constant, and the other was varied, the same RT and MT differences were found that have been indicated in earlier studies. Workload ratings were similar for all of the PM tasks, with the exception that PME was rated as having less workload than PMH.

<u>Math Functions</u>

A significant difference in RSTs was found due to the complexity of the different mathematics tasks, following the expected trend: the RSTs were shortest for the G/L task, followed by the Wittenborn task (W), and the EQ task, $F(2, 8) = 24.00$, $p<.01$. There was a significant effect of task on percent correct as well, $F(2, 16) = 16.5$, $p<.01$.

Response selection times for the Math + TA condition were slightly faster than for the math tasks alone. This could be an effect of training, since all of the TA tasks were presented after the response-selection-only tasks. Two other findings were of interest: There was a significant interaction between task and right/left responses, $F(2, 16) = 10.69$, $p<.01$. Right RSTs were faster than left RSTs, $F(2, 8) = 10.73$, $p<.01$, due primarily to the EQ task, in which left movements (less) were twice as slow as right movements (greater). Also, an effect was found for task on MT, $F(2, 16) = 6.31$, $p<.01$; however, since the conditions having the most control reversals also had the longest MTs, the extra time taken by the reversals accounts for this effect.

Workload ratings mirrored task complexity, with EQ > W > G/L, $F(2,8) = 21.11$, $p<.01$ (Figure 4). Single (RT only) task workload was not significantly different than dual task workload.

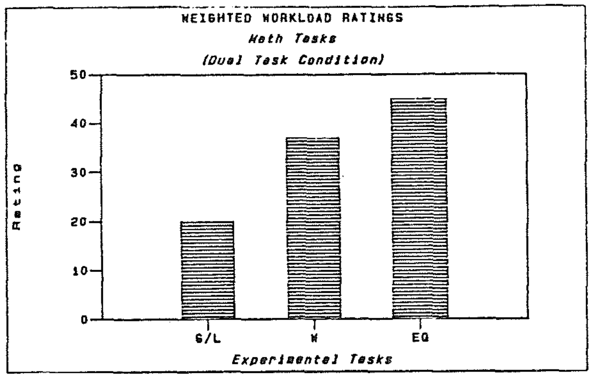

Figure 4. Weighted workload ratings for math tasks.

<u>Time Estimation</u>

In the time estimation tasks, no effects were found for percent correct, number of reversals, or MTs. Left (5-sec) and right (10-sec) responses were examined separately. For left responses: time estimates were significantly longer for the TS task than for T, in both the single and dual task modes, $F(1, 8) = 11.46$, $p<.01$. Estimates in both TS and T were also longer in the single task condition than in the dual task condition, $F(1, 8) = 8.41$, $p<.05$.

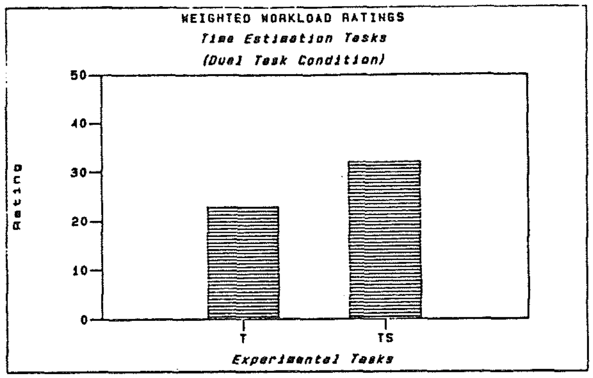

Figure 5. Weighted workload ratings for time estimation tasks.

Right (10 sec) responses showed somewhat similar results. Estimations were longer in the single-task condition for both TS and T, but there was no difference in estimates between the T and TS tasks. Overall, in the time estimation tasks, the 5-sec estimations were more accurate than 10-sec estimations, which were generally too short.

Workload ratings ranked TS as harder than T, $F(1, 8) = 7.2$, $p<.05$, and showed no difference between the single and dual task conditions (Figure 5). A somewhat surprising finding was that many subjects considered the TS task, which involved estimating time as well as solving an equation, to be easier than the EQ task. Reportedly, this was because they did not feel as much time pressure in solving the equation, since the solution to the equation could be completed at any time up to the end of the shortest of the two estimation intervals.

<u>Rhyme Tasks</u>

The delayed rhyme task (SRYM) resulted in significantly faster RSTs, $F(1, 8) = 25.35$, $p<.01$ (Figure 6), and a greater percent correct (49% vs. 47%), $F(1, 8) = 13.3$, $p<.01$, than the immediate rhyme task. No difference was found between single and dual task conditions in RSTs; however, there was a significant difference in percent correct in favor of the dual task condition, $F(1, 8) = 8.4$, $p<.05$, probably due to training. No differences for MT or reversals were found. "No" (left movement) responses in both tasks were much slower than "yes" (right movement) responses, $F(1, 8) = 19.01$, $p<.01$, a common finding.

There were several interactions: the difference in RST between RYM and SRYM decreased with training, as illustrated by a Task x Condition (RYM vs. SRYM x RS vs. FB) interaction, $F(1, 8) = 12.01$, $p<.01$; the "yes/no" effect was more pronounced in the RYM task than in the SRYM task, as shown by a Right/Left x Task interaction, $F(1, 8) = 13.78$, $p<.01$; and practice reduced this "yes/no" effect, illustrated by a Right/Left x Condition interaction $F(1, 8) = 7.78$, $p<.05$.

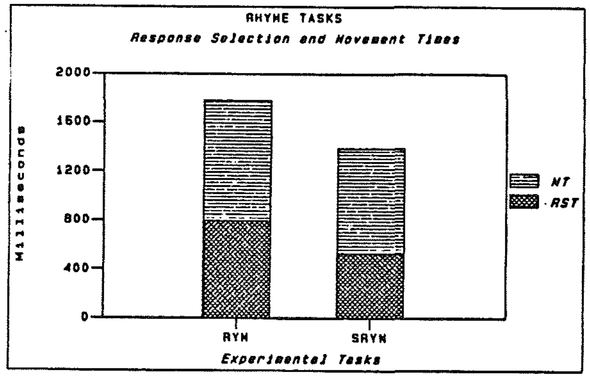

Figure 6. Response selection and movement times for rhyme tasks.

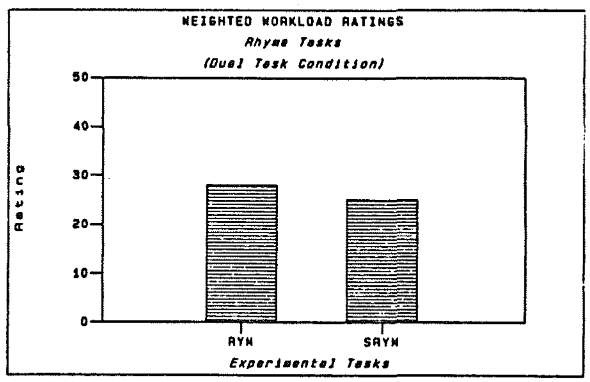

Figure 7. Weighted workload ratings for rhyme tasks.

The RYM task was rated as having greater workload than the SRYM task, $F(1, 8) = 6.65$, $p<.05$ (Figure 7) both in RT and RT/Fitts TA conditions. Reasons for this are discussed in the following section.

DISCUSSION

In general, task complexity had the predicted effect on response latency; that is, the more complex the required cognitive processing, the longer it took before a response was selected. The additional processing demanded by the response execution task, however, was not reflected in RST. In fact, for some tasks (e.g. math tasks), RSTs in the dual-task modes were somewhat faster than in the response-selection-only mode. The probable cause for this counterintuitive result could be training; by the time subjects began performing the dual condition, they were familiar with all of the cognitive tasks, and, since they had previously participated in a Fittsberg study, were practiced in target acquisition.

Workload ratings also reflected task complexity, with a few unforeseen results. Several of the response-selection-only tasks were rated as having somewhat higher workload than the same task in the Fitts TA mode. Since all of these tasks were performed first, this again could be the result of training. The simultaneous rhyme task (RYM) was seen as being more loading than the Sternberg rhyme task (SRYM), even though SRYM involved no memory and used the same type of words. The equation task (EQ) was perceived as being much more difficult than the combination of equation and time estimation (TS), even though the latter involved an additional processing step. The apparent reduction in time pressure mentioned earlier seems to have been an overriding factor here. The equation task took the longest to complete; thus, removing the pressure of having to do it immediately served to greatly reduce its perceived workload (TS was one of the tasks rated lowest in workload, even though its EQ task component was rated as highest).

Performance

Since MTs were not affected by the complexity of the response selection task, it is reasonable to assume that any decision making was completed prior to the movement phase--or, at least, that whatever processing did carry over was sufficiently minimal to be accomplished simultaneously with movement, causing no detriment to MT.

Task complexity did have an observable effect on percent correct; it was largest with the easiest tasks (96%), and smallest with the most difficult tasks (82%). Control reversals did not follow the same pattern, as there were relatively high numbers of reversals for some of the less complex tasks (O/E, RYM). One possible explanation for this is that these tasks were so simple that they were performed "enroute"; that is, subjects may have "jumped the gun" by starting movement in one direction before they had completely processed the stimulus, then finished processing, changing direction if necessary once the stimulus was fully absorbed.

The results of this study indicate that information processing of the response selection task and of the target are done concurrently, since dual task RSTs were, in general, equal to or only slightly greater than the single task RSTs. This is in keeping with previous findings.

Workload

Workload ratings for the response selection tasks paralleled almost exactly response latencies, especially at the extremes: for the immediate response tasks, G/L was considered to be the easiest task (WL = 22) and resulted in the shortest mean RST; EQ was considered to be most loading (WL = 47), and had the longest mean RSTs. The time estimation task (T), in which there was no pressure for a fast response, was also considered to be very low in workload (WL = 23). This indicates that subjects were very sensitive to the relative amounts of required processing and in their perception of the time pressure imposed on them. These were each reflected in their evaluations of the tasks.

Dual task workload was not consistently greater than the same task presented in the single-task mode. This replicated, in general, findings of earlier studies (e.g., Hart, Shively, Vidulich, & Miller, 1985). A tentative explanation for this would, again, be training effects, negating the perception of additional load. Subjects had had enough practice on the basic TA task and the single-task response-selection tasks that the combined task might have imposed no extra load. Another possibility is that most of the perceived workload was in the response selection phase; therefore, the

Fitts TA was experienced as an equivalent task, even though RST's indicated that additional processing was required. Since their existed a functional relationship between the response selection task and the target acquisition task, the latter may have been viewed as merely an extension of the former.

With regard to various specific tasks, the type of memory involved did not appear to have as much impact on workload and RST as did the specific design of the task. That is, the concurrent memory tasks were not, as a group, faster or slower than the recent memory or long-term memory tasks, with some interesting anamolies. For example, examining the RYM and SRYM tasks, it would seem logical that the concurrent processing task, RYM, would have been at least as easy, if not easier, than a long-term memory task; however, the immediate comparison (RYM) resulted in longer RSTs, more errors, and higher workload ratings that SRYM. A factor that may have contributed to this was that, in SRYM, the same word was compared with each other word continuously through the block of trials; in RYM, however, two completely different words were presented on each trial.

The major direction of movement differences were found for tasks in which right or left signified a yes/no response. The lag in RST for a "no" response is of consequence in the real-world cockpit environment in that the discovery that instrument readings (e.g., altitude, heading, fuel supply) are not as they are supposed to be usually signifies trouble--and this may be a situation that calls for the quickest possible action. Also of interest was the fact that the three tasks with the longest RT's--EQ, W, and SET--all involved dealing with numbers. The solution of a simple function in EQ took, on the average, one minute longer than the next slowest task and resulted in more mistakes. This has important operational implications as well.

There were many incorrect responses for the SET and EQ tasks. The SET task is similar to those performed in flight; headings, altitudes, radio frequencies (i.e., sets of numbers), are continually being updated, and the operator is often required to compare current sets of values to previous sets. The design of SET made this activity particularly difficult because subjects could not "chunk" the three numbers; each digit had to be tested against the previous values, and remembered individually.

A key issue in these findings is the difference between the actual Fittsberg RST and WL ratings that were observed in the present study, and what might be predicted on the basis of simply adding the levels of the two component tasks. If RST and WL are cumulative, that is, if each additional task imposes its own requirements on top of those of the previous task, then one would predict that RST(FB) = RST + RT(TA) and WL(FB) = WL(RST) + WL(TA). Table 2 illustrates the RST and WL that would be expected if this were the case. However, the actual figures are much less than this sum; in fact, in some cases, the obtained RST or WL was equal to or only slightly greater than that of either the response selection or response execution task alone.

Table 2

Predicted versus Observed WL and RST

TASK	RS	RE	RST SUM	OBS	RATIO	RS	RE	WL SUM	OBS	RATIO
PM	.479	.45	.929	.456	.49	31	20	51	22	.43
G/L	.423	.45	.873	.454	.52	24	20	44	20	.45
O/E	.485	.45	.935	.495	.53	26	20	46	22	.48
RYM	.803	.45	1.253	.729	.58	35	20	55	28	.51
SRYM	.528	.45	.978	.553	.56	25	20	46	25	.54
SET	.910	.45	1.360	.817	.60	39	20	59	34	.58
T	7.413	.45	7.863	6.612	.84	19	20	39	23	.59
W	.932	.45	1.382	.838	.61	42	20	62	37	.60
EQ	2.016	.45	2.466	1.744	.71	49	20	69	45	.65
TS	7.837	.45	8.287	7.087	.85	28	20	48	32	.67

In this study, the response selection tasks that required only one processing step (e.g., PM, O/E, G/L) were most easily integrated with the TA task, and evidenced the largest discrepancy between the additive prediction of RST and WL and the actual figures. The cognitive processing of these tasks was simple enough to be accomplished in parallel with TA, without additional cost; and dual task WL ratings and RSTs are essentially equivalent to those of the response selection tasks alone. In keeping with this, WL ratings for all of the dual tasks were found to be highly correlated with RST. The processing tasks requiring more than one step (e.g., W required addition + comparison; SET required memory + comparison; EQ required arithmetic problem solving + memory + comparison) were less easily integrated, and the observed WL and RST in these tasks came much closer to the additive predictions. If the tasks were not at all functionally related, the expected ratio of observed to predicted WL would be >1.

Perceived time pressure, rather than experimental manipulation of time pressure, contributed significantly to rated workload, with unforeseen results. For example, the EQ task was rated as having the most workload, and resulted in the largest number of errors and the longest RSTs. However, the TS task (which contained the same equations with the additional task of time estimation), was rated as one of the easiest tasks and resulted in minimal errors - because subjects were able to perform the mental arithmetic calculations at their leisure during the time estimation interval. Since the TS task was a combination of two cognitive tasks, time estimation and arithmetic problem solving, TA actually imposed a third requirement. The predicted WL for TS in the "dual" task condition would be 19(T) + 49(EQ) + 20(TA) = 88. The obtained WL rating for this task, however, was 32 - less than half the prediction. This would seem to indicate that reducing or removing the significant elements contributing to WL, as well as increasing the functional relatedness of tasks, can greatly reduce experienced workload.

The results of this study have implications for laboratory as well as operational tasks. In functionally related tasks, processing for response selection and execution appear to be done in tandem. The cognitive complexity of the task profoundly affects the response selection part of the task, but only the physical properties of the target affect the difficulty of its acquisition. Subjects can measurably differentiate the cognitive complexity of tasks - both in terms of performance (actual motor responses) and in terms of perceived workload. Also, the more functional overlap that exists among tasks that are to be performed concurrently or serially, the more the operator can mentally integrate the tasks, and the less the cost in terms of performance and experienced load.

In view or this, human factors engineers must concentrate on keeping cognitive complexity to a level that is manageable and has acceptable consequences in terms of response latencies. Additionally, since the cost of imposing more tasks can vary widely, the nature and relatedness of the simultaneous or serial tasks required of the human operator must be taken into account.

Indications were present on some tasks that training can have the effect of not only improving performance, which is intuitively predictable (as shown in the math tasks, which had the longest RTs, and possibly the most room for improvement); but can also function to reduce perceptions of workload in an equivalent or objectively more difficult task. This was illustrated by the several tasks in which the single task, presented first, was rated as being higher in workload than the same task in the dual condition. One of the possible effects of training is to facilitate integration of the tasks being performed. Therefore, training apparently can, to a certain extent, compensate for increased task loading.

REFERENCES

Fitts, P. M., & Peterson, J. R. Information capacity of discrete motor responses. Journal of Experimental Psychology, 1954, 47, 381-391.

Hart, S. G., Battists, V., & Lester, P. T. POPCORN: A supervisory control simulation for workload and performance research. Proceedings of the 20th Annual Conference on Manual Control, Sunnyvale, CA., 1984.

Hart, S. G., Sellers, J. J., & Guthart, G. The impact of response selection and response execution difficulty on the subjective experience of workload. Proceedings of the 28th Annual Meeting of the Human Factors Society, San Antonio, TX., 1984.

Hart, S. G., Shively, R. J., Vidulich, M. A., & Miller, R. C. The effects of stimulus modality and task integrality: Predicting dual-task performance and workload from single task levels. Paper presented at the 21st Annual Conference on Manual Control, Columbus, OH., 1985.

Hartzell, E. J., Gopher, D., Hart, S. G., Dunbar, S. L., and Lee, E.

Moray, N. Attention: Selective Processes in Vision and Hearing. New York: Academic Press, 1970.

Polich, J., McCarthy, G., Wang, W. S., & Donchin, E. When words collide: Orthographic and phonological interference during work processing. Biological Psychology, 16, 1983, 155-180.

Reid, G. B., Shingledecker, C. A., Nygren, T. E., & Eggemeier, F. T. Development of multidimensional subjective measures of workload. IEEE Proceedings of the International Conference on Cybernetics and Society, 1981, 403-406.

Staveland, L., Hart, S. G., & Yeh, Yei-Yu. Memory and subjective workload assessment. Paper presented at the 21st Annual Conference on Manual Control, Columbus, OH, 1985.

Sternberg, S. Memory scanning: New findings and current controversies. Quarterly Journal of Experimental Psychology, 27, 1-32.

Wierwille, W. W., and Casali, J. G. A validated rating scale for global mental workload measurement applications. Proceedings of the 27th Annual Meeting of the Human Factors Society, Norfolk, VA, 129-133.