Acta Psychologica 71 (1989) 53–87
North-Holland

WORKING MEMORY IN THE ACQUISITION OF COMPLEX COGNITIVE SKILLS *

Robert LOGIE

University of Aberdeen, UK

Alan BADDELEY

MRC Applied Psychology Unit, Cambridge, UK

Amir MANÉ, Emanuel DONCHIN and Russell SHEPTAK

University of Illinois at Urbana-Champaign, USA

This paper reports three experiments using the secondary task methodology of working memory, in the task analysis of a complex computer game, 'SPACE FORTRESS'. Unlike traditional studies of working memory, the primary task relies on perceptual-motor skills and accurate timing of responses as well as short- and long-term strategic decisions. In experiment 1, highly trained game performance was affected by the requirement to generate concurrent, paced responses and by concurrent loads on working memory, but not by the requirement to produce a vocal or a tapping response to a secondary stimulus. In experiment 2, expert performance was substantially affected by secondary tasks which had high visuo-spatial or verbal cognitive processing loads, but was not contingent upon the nature (verbal or visuo-spatial) of the processing requirement. In experiment 3, subjects were tested on dual-task performance after only 3 hours practice on Space Fortress, and again after a further five hours practice on the game. Early in training, paced generation of responses had very little effect on game performance. Game performance was affected by general working memory load, but an analysis of component measures showed that a wider range and rather different aspects of performance were disrupted by a visuo-spatial memory load than were affected by a secondary verbal load. With further training this pattern changed such that the differential nature of the disruption by a secondary visuo-spatial task was much reduced. Also, paced generation of responses had a small effect on game performance. However the disruption was not as dramatic as that shown for expert players. Subjective ratings of task difficulty were poor predictors of performance in all of the three experiments. These results suggested that general working memory load was an important aspect of performance at all levels

* This work was funded by the Defense Advanced Research Projects Agency (DARPA), under contract MDA903-84-K-0065, and subcontracted through the University of Illinois to the Medical Research Council (UK). We are grateful to DARPA for its support. We are also grateful to the technical staff at the University of Illinois at Urbana-Champaign, and in particular, to W. Sherman, who was largely responsible for the computer programming involved.

Requests for reprints should be sent to R. Logie, Dept. of Psychology, University of Aberdeen, Aberdeen, UK.

of training. The greater disruption by paced responses in experts was interpreted as suggesting that response timing is important for expert performance. The change with training in the differential interference from a visuo-spatial versus a verbal secondary task was interpreted as suggesting that perceptual-motor tracking control is an important and demanding aspect of novice performance but that it is a highly automated skill in the performance of experts. The implications of this methodology for the study of cognitive workload, of skill acquisition and the potential implications for the theoretical development of working memory are discussed.

Introduction

A theme underlying the approach to the Space Fortress project is that performance on the game may be fruitfully subdivided into a number of subcomponent skills. Our approach aimed to test this directly by means of a secondary task procedure which has proved fruitful in the development of the concept of working memory (Baddeley and Hitch 1974; Baddeley 1986). The main approach of this concept has been as a functional analysis of temporary storage and manipulation during information processing. It assumes a central executive for complex decision and control processes, and a number of subsidiary slave systems, thought to be involved in specific processing.

A major contribution of this approach has been in providing a coherent, functional description of short-term verbal storage and manipulation, within the context of an articulatory loop: a 'slave system' thought to act as a subvocal rehearsal buffer (e.g. Baddeley et al. 1975a; Salamé and Baddeley 1982; Vallar and Baddeley 1982). A second slave system, the Visuo-Spatial Sketch Pad, is thought to provide similar functions for visuo-spatial material (Baddeley et al. 1975b; Baddeley and Lieberman 1980: Logie 1986; Logie and Baddeley 1989).

The potential scope of this concept is, in principle, sufficiently wide to provide insight into an enormous number of everyday tasks of widely varying complexity. Some studies have shown working memory to be useful in the study of reading (Baddeley 1979; Baddeley et al. 1985), and counting (Hitch 1978; Logie and Baddeley 1987). However, most of the tasks so far chosen for study have been relatively simple or have had strong cognitive components, with relatively little perceptual-motor or speed content.

In contrast, Space Fortress relies on both perceptual-motor skills and more cognitive skills such as those involved in both moment to moment decisions and in more long-term strategies. The experiments described had two aims: first to provide an analysis of a complex task in terms of subcomponent skills, and second, to investigate the use of a complex task environment as a test of the generality of the working memory framework and methodology.

The task analysis of this game was attempted using dual-task techniques, typically employed in experiments within the area of working memory. The procedure rests on the assumption that where mutual interference occurs in performing two tasks simultaneously when compared to performing each task alone, that the tasks share some cognitive resource. The secondary tasks are chosen to involve specific resources and therefore, any mutual disruption between the game and a particular secondary task would suggest that the cognitive functions important for the secondary task are also involved in the game.

This assumption is not entirely without polemic (e.g. Allport 1980; Navon 1987; Wickens 1980). Nonetheless it has been particularly fruitful in providing coherent sets of converging evidence for the various components of working memory (Baddeley 1986).

Our experiments were concerned with four questions, first the role of response competition in performance, second the importance of general working memory load and third, the relationship between visuo-spatial processing and perceptual-motor control in the game. Finally, the experiments studied the role of these factors at various levels of expertise on the game.

Experiment 1

Method

Secondary task procedures

The secondary tasks for experiment 1 were chosen to cover a wide range of cognitive and perceptual-motor functions that might be involved in the game, including response factors, timing, and memory load. Note that the secondary tasks were not given to subjects in the order shown. To provide descriptive coherence, the tasks have been arranged in three sections, appropriate to assumed sets of component skills.

The intention of the first three secondary tasks was to investigate interference at the response level between game performance and various secondary tasks. These involved auditory presentation of a random sequence of days of the week, presented at a rate of one day every three seconds, through a Votrax voice synthesizer. The conditions varied in the response that the subject was required to make.

(1) Repeat a day. The subject's task was to repeat the day presented as quickly as possible. Response times and omitted responses were recorded by the computer using a voice key, while errors (responding with the wrong day) were recorded by the experimenter. When combined with Space Fortress, the stimuli were not presented during explosions, but were otherwise continuous throughout.

(2) Repeat next day. The procedure for this task was virtually identical to that for repeat a day, except that subjects had to repeat the day of the week that immediately followed the day presented.

(3) Tap to a day. In this task, the stimuli were as for the previous two conditions, however, subjects were to tap a foot pedal as quickly as possible after presentation of a day. Response times and omitted responses were recorded by the computer.

The next set of tasks involved the generation of a response at a pre-set rate, but with no stimulus present. The intention with these tasks to study the effect on those aspects of the game that require timing of responses. The conditions vary in the rate and type of responses required.

(4) Slow tapping. In this condition, subjects were to tap a foot pedal once every three seconds. Prior to each block, the rate was demonstrated for 20 seconds, with subjects tapping in response to the tick of a metronome. Inter-tap intervals were recorded by the computer.

(5) Fast tapping. The procedure here was similar to that for slow tapping, except that subjects were to alternate between two foot pedals at a two per second rate.

(6a) Repeated articulation. In this condition, subjects were to repeat a single word ('GO') at a two per second rate. As for tapping, the rate was demonstrated prior to each trial block. The inter-response intervals were recorded by the computer via a voice key.

The last set of tasks involved an investigation of the working memory load required by the game.

(6b) Articulatory suppression. This was identical to condition 6a, considered in the context of its effect on memory performance, in particular, recall of which letters indicated a 'foe' mine. Numerous previous studies have shown repeated articulation of an irrelevant word to be associated with specific impairment of short-term verbal memory (Baddeley 1986).

(7) Word span. In this condition, the experimenter read to the subject lists of single syllable, high frequency words for immediate ordered recall. During a first testing session, each subject's word span was determined using a standard memory span procedure. Subjects were initially given three lists of two words for recall. List length was increased by one word after every group of three lists until the subject was unable to recall accurately two of the three lists at a given level. For each subject, we calculated the mean number of words from the three longest lists that were accurately recalled. Span for that subject was taken as the nearest whole number equal to or below the calculated mean. Collection of data on this task alone and when combined with Space Fortress, involved presenting word lists with the number of words constant and equal to the span of the individual subject. As with the previous conditions, word lists were not presented during explosions, although subjects could recall during these periods.

(8) Sentence span. This condition was based on a procedure reported by Daneman and Carpenter (1980) and was similar to word span except that subjects were given sets of simple sentences, such as 'THE BUTCHER ATE THE APPLE; THE ARTIST CUT THE GRASS'. The experimenter then read either 'THE OBJECTS' or 'THE SUBJECTS'. The task was to recall the objects or the subjects of the sentences as appropriate, in the order in which they were presented. In the example given, if 'THE OBJECTS' were required, the subject should respond 'APPLE, GRASS'. Span for this material was measured for each subject, using the procedure for word span, except that four groups of sentences were given at each group length, starting with two sentences. Procedure otherwise was as for word span.

General procedure

Subjects were tested over three sessions, with a gap of at least one day between sessions. The first session was to provide practice and to collect control data on the game and on each of the secondary tasks alone. During this first session, the word span and sentence span were determined, with no further practice on these tasks. For all remaining secondary tasks, subjects were given practice on each task, followed in each case, by five minutes on the secondary task alone to provide control data. The conditions were given in the order; Repeat-a-day, Tap-to-a-day, Repeat-next-day, Slow tapping, Game alone, Word span, Articulatory suppression, Sentence span, Rapid tapping.

In the two subsequent sessions, subjects were given five minutes on each span task alone, prior to combining with the game. Two minutes of control data were collected for each of the remaining secondary tasks, followed by the appropriate dual task condition. Subjects played the game alone once at the start of each session. In the second session, the order of conditions was otherwise as for session one, with each secondary task performed alone prior to being combined with the game. In session three, the order of the experimental conditions was reversed, however, performance of the secondary task alone always preceded the dual task.

After every game, the subjects were asked to recall the letters that had been assigned to foe mines for that game. Finally, after each condition, subjects were asked to rate the difficulty of the task just completed. They were given a ten-point scale, where a rating of one indicated that the task was extremely easy with virtually no effort required, and a rating of ten indicated that the task was virtually impossible to carry out.

Subjects

At this stage, we were particularly interested in an analysis of expert performance on the game. Therefore we chose six subjects who were given a minimum of twenty hours practice on the game, prior to the introduction of any secondary tasks. Other investigators who have used Space Fortress, have found that ten hours practice is sufficient to gain relatively skilled performance on the task, and the learning curves of the subjects chosen for these experiments suggested that they were approaching asymptotic performance. Subjects were all right-handed males, aged 18–20 years, and students at the University of Illinois at Urbana-Champaign, where the experiments were conducted. Subjects were paid for their participation. Session one lasted approximately 90 minutes. Sessions two and three lasted about two hours each. In each case, subjects were given a short break about half way through the session.

Results

The game provides two classes of performance measure: a total score, and a large number of measures reflecting the microstructure of performance. Some of the component measures contribute to the total score. Some could also be reasonably used as measures of components of performance on the game. We chose seventeen of these measures for our detailed analysis and a list of these is given in the appendix.

The intention at the first stage of analysis, was to look for general indications as to patterns of impairment in game performance. As such, the data were first examined in terms of the overall, relative effects on the gross measure of game score and the effects of game playing on secondary task performance. The more detailed measures of performance were considered in a second set of analyses of effects on game microstructure.

Game scores for game alone and with each of the secondary tasks were entered into a one-way analysis of variance, investigating the effects of secondary tasks (9 levels). The analysis confirmed that the presence of secondary tasks resulted in significant impairments in game score ($F(1,8) = 8.668$; $MSe = 371024$; $p < 0.001$). We then considered the effects of particular secondary tasks on game score, by means of multiple pairwise comparisons using the Newman-Keuls test (Keppel 1973). We also analysed performance on each secondary task according to whether it was performed alone or in combination with the game, and the results of these analyses are reported. Clearly since different tasks are involved, a formal comparison between secondary tasks would be rather difficult to interpret, even where the measures involved are apparently similar. The analyses have been arranged according to our grouping of secondary tasks.

Response to an external stimulus

Table 1 shows the effects on game score of a timed response to a secondary stimulus. Of the three types of response, only tap-to-a-day produced a significant deterioration in overall game score. Table 1 also shows the secondary task performance with and without the game. Response times to the stimuli were compared using individual t-tests (2-tail). All three tasks showed some slowing of response time to an external stimulus when combined with playing Space Fortress. For tapping to a day, response times were on average, 405 msec slower when both tasks were required ($t = 3.64$ $df = 28$; $p < 0.01$). They were also more variable ($t = 8.69$, $df = 28$; $p < 0.01$), as measured by the mean difference in standard deviation.

Repeating the same day was slower by 301 msec ($t = 2.89$, $df = 28$; $p < 0.01$), and more variable ($t = 8.95$, $df = 28$; $p < 0.01$), and responding with the next day was slower by 230 msec ($t = 1.88$, $df = 28$: $p < 0.05$), as well as significantly more variable, ($t = 4.54$, $df = 28$; $p < 0.01$) suggesting that these tasks resulted in some impairment in overall performance, but that subjects maintained their level of game performance.

Table 1
Mean game score, response time (msec), mean difference in RT standard deviation (*SD*) for game alone, and with tap to a day, repeat a day and next day as secondary responses.

	Control	Tap day	Repeat day	Next day
Game score	5696	5059 [b]	5700	5585
Sec. task alone (msec)		611	807	1058
Sec. task + game (msec)		1016	1108	1288
Difference (msec)		405 [b]	301 [b]	230 [a]
Mean difference in *SD* (msec)		236 [b]	188 [b]	142 [b]

[a] $p < 0.05$; [b] $p < 0.01$.

Generation of responses

Table 2 shows the summary data for slow and rapid tapping, and for repeated articulation. Game score was significantly impaired by slow tapping (CR = 6; $p < 0.01$), and by rapid tapping ($CR = 7.626$; $p < 0.01$), resulting in a drop of over 1,000 points in each case. There was no impairment with concurrent articulation.

Performance on the secondary task appears to have been affected by combination with the game, resulting in an increased rate of responding. However, only slow tapping was affected significantly ($t = 1.94$, $df = 28$; $p < 0.05$), and there was no significant change in the variability of inter-response intervals for any of these tasks.

Table 2
Mean game score, and inter-response intervals (msec) for game alone, and for slow tapping, rapid tapping and repeated articulation as secondary tasks.

	Control	Slow tap	Rapid tap	Articulation
Game score	5696	4641 [b]	4355 [b]	5103
Sec. task alone (msec)		3035	643	441
Sec. task + game (msec)		2240	529	396
Difference (msec)		795 [a]	114	45
Mean difference in *SD*		38.9	25.8	7.4

[a] $p < 0.05$; [b] $p < 0.01$.

Table 3
Mean game score, inter-response interval (msec), and span data for game performance and repeated articulation, word span and sentence span as secondary tasks.

	Control	Articulation	Word span	Sentence span
Game score	5696	5103	4634 [a]	4920 [a]
Sec. task alone (msec and span)		441	5.03	2.43
Sec. task + game (msec and span)		396	4.30	2.26
Difference		45	0.73 [a]	0.17

[a] $p < 0.01$.

Working memory load

Summary data for game performance and performance on word span and sentence span are shown in table 3. Results for concurrent articulation are repeated in this table for comparison. Word span produced a significant degree of impairment in game score ($CR = 6.040$; $p < 0.01$), as did sentence span ($CR = 4.413$; $p < 0.05$). There was a tendency for both sentence span and word span to be reduced when performed with the game, however only word span showed a significant decrement ($t = 2.789$, $df = 22$; $p < 0.01$).

Discussion

Taken as a whole, it is clear that with the exception of articulatory suppression, all of our secondary tasks resulted in some impairment in performance when combined with the game, although with repeat-a-day, tap-to-a-day and next-day, this impairment appeared in the secondary task. However, the relative effects differed with the particular secondary task chosen. Table 4 summarises all of the mean game scores across all secondary tasks, along with the mean rated difficulty of each task alone and with the game.

Initial inspection suggests that the relative impairments in game performance could not be predicted simply on the basis of subjective levels of difficulty. On the basis of a model of cognition that suggested the use of general purpose resources, it would be reasonable to expect that the secondary tasks rated as most difficult when performed singly, should also result in the largest performance decrements in the dual task conditions. This is clearly not the case. For example, the task which appears to produce the greatest overall decrement in game score, rapid tapping, was rated as one of the easiest tasks when performed alone. The low, non-significant correlation between these measures (0.354), supports this conclusion.

The picture is similar if we consider decrements in secondary task performance. Repeat-a-day and tap-to-a-day were rated as being the easiest of any of the tasks when

Table 4
Percentage decrements in game score and secondary task performance and mean rated difficulty for all conditions in experiment 1.

	Decrement in game score (%)	Rated difficulty game + task	Decrement in sec. task (%)	Rated difficulty task alone
Repeat day	0	4.42 [c]	37.30 [b]	1.39 [b,d]
Game alone		3.72		
Next day	1.95	5.08 [a]	21.7 [a]	2.17 [b]
Repeated articulation	10.41	5.67 [b]	10.2	3.33 [b]
Tap-to-day	11.18 [b]	5.08 [a]	66.30 [b]	1.39 [b]
Sentence span	13.62 [b]	5.83 [b]	7.0	4.67
Slow tap	18.52 [b]	6.83 [b]	26.2 [a]	2.56 [b]
Word span	18.64 [b]	6.67 [b]	14.5 [b]	4.17 [b]
Rapid tap	23.54 [b]	5.50 [b]	17.7	2.00 [b]

[a] $p < 0.05$; [b] $p < 0.01$.
[c] Significance levels are for comparison with rating for game alone.
[d] Significance levels are for comparison with rating for secondary task with the game.

performed alone, but these tasks showed the largest percentage impairments when performed with the game. The correlation in this case is significantly negative ($r = -0.747$; $p < 0.05$), the opposite to what would be predicted by a simple general resources model. However, relative percentage decrements in secondary task performance should be treated with caution since they reflect changes in a range of rather different tasks.

These results thus suggest that impairments due to a secondary task are not simply due to general cognitive load, and that the effects may reflect interference with specific components of the primary task. We examined this notion further in an analysis of game microstructure.

Effects on game components

We next considered the effects of the presence of our secondary tasks on the various component measures listed in the appendix. The scores for each game component for game alone and with each secondary task were entered into separate analyses of variance, and the relative effects of each secondary task on each game component were examined using Newman-Keuls multiple comparison procedure. We were primarily interested in the effect of specific secondary tasks on each game component. In view of the number of components involved, and for the sake of clarity, the results of the analyses of variance will not be reported. Instead we will report the results of the post hoc comparisons using Newman-Keuls tests (Keppel 1973) to show the pattern of disruption by individual secondary tasks. The post hoc comparison for any one game component was carried out only if the overall analyses of variance for that component was significant at the 1% level or better.

Table 5 shows the mean scores for each game component for game alone and with each of the secondary tasks. Table 6 summarises the results of multiple comparisons, in

Table 5
Mean scores on component measures for game alone and with each secondary task in experiment 1.

Game component	Control	Repeat day	Tap day	Next day	Slow tap	Fast tap	Articulatory suppression	Word span	Sentence span
MOVMNT	879	838	885	865	945	885	900	896	938
NBDBLE	4.33	5.75	3.33	6.58	2.67	5.42	8.42	3.50	3.00
NBDIFF	0.278	0.167	0.333	0.417	0.333	0.833	0.917	0.333	0.250
NBDINT	0.61	1.33	1.67	2.00	2.75	2.25	4.42	2.25	2.92
NFODES	33.8	34.8	31.3	33.8	28.1	28.6	31.8	30.5	26.3
NFOHITS	393	413	367	409	332	355	393	367	312
MNOVIN	4.28	3.17	3.83	2.92	6.83	4.33	4.17	4.42	8.25
NSHDMG	3.56	4.50	5.92	4.25	6.33	6.42	5.83	6.33	7.75
NWRAP	0.39	0.67	0.33	0.00	1.58	0.92	1.50	0.25	1.67
PCNTBON	93.9	92.5	92.5	92.5	90.8	80.8	88.3	83.3	92.5
RTIFF	89.7	90.7	95.1	103.6	99.8	112.1	96.2	111.7	100.6
SHTEFF	89.3	89.0	88.3	88.3	84.0	87.1	85.8	87.5	81.3
TIMKLFO	36.9	37.2	38.2	38.2	41.3	40.6	40.0	40.4	40.1
TIMKLFR	29.3	30.75	33.0	31.3	31.8	33.0	31.1	32.5	32.6
TOTSHOTS	492	506	454	506	437	448	500	458	427

Table 6
Effects of secondary tasks on component measures in experiment 1.

Secondary task	Components affected
Repeat-a-day	None
Tap to day	None
Next day	None
Slow tap	NFOHITS, NFODES, SHTEFF
Rapid tapping	NFOHITS, NFODES, SHTEFF
Repeated articulation	NBDINT
Word span	NFODES, RTIFF
Sentence span	RTIFF, TOTSHOTS

terms of the groups of component measures affected by particular secondary tasks. Given the relatively large number of comparisons involved, we set a conservative acceptable significance level for these analyses ($p < 0.01$).

From the table, it appears that the overall effects on game score did not reflect a general load on processing resources, but rather that the secondary tasks affected different aspects of game performance. Slow tapping affected the number of times that the fortress was hit by a missile or by a mine (NFOHITS: $CR = 6.099$; $p < 0.01$), the number of times the fortress was destroyed (NFODES: $CR = 5.953$; $p < 0.01$), and the proportion of missiles fired by the ship that hit an appropriate target (SHTEFF: $CR = 7.378$; $p < 0.01$). Rapid tapping affected the same variables as slow tapping (NFOHITS: $CR = 8.108$; $p < 0.01$, NFODES: $CR = 7.847$; $p < 0.01$, and SHTEFF: $CR = 11.185$; $p < 0.01$). Repeated articulation affected only one variable; the frequency with which the IFF button was pressed with an inter-press interval outside the accepted range (NBDINT: CR = 6.750: $p < 0.01$). In contrast, word span and sentence span both affected the delay in responding to the appearance of a foe mine (RTIFF: $CR = 5.960$ and $CR = 5.854$; $p < 0.01$). In addition, word span affected the number of fortress destructions (NFODES: $CR = 5.435$; $p < 0.01$). while sentence span affected the total number of shots fired by the ship (TOTSHOTS: $CR = 5.116$; $p < 0.01$).

The number of fortress destructions (NFODES) can arise from a variety of factors, depending on the efficiency of using the missiles, pressing the IFF button correctly, pressing the firing button twice at the correct time, and so on. In this sense, it can be taken as a relatively gross measure of performance. However, shooting efficiency (SHTEFF) and number of hits on the fortress (NFOHITS) could both the thought of as measures of aiming and timing of responses to events in the game. These appeared largely unaffected by the requirement to respond to secondary stimuli, but were affected by the requirement to generate responses at a set rate. A further measure of timing, the number of bad intervals (NBDINT) was also affected by one of these secondary tasks, namely repeated articulation. In this respect, concurrent articulation appears to produce impairment due to the timing and generation required, rather than any effect on short-term verbal memory load.

Subjects could accurately recall the foe letters at the end of each game, regardless of whether there was a secondary task. This suggests that foe letter recall was a relatively insensitive measure, perhaps due to the minimal memory load imposed by remembering a set of three letters. A more sensitive measure might be the time for recall during the game, rather than accuracy after the game. This appeared to be the case, since the main effects of memory load arose in the time taken by subjects to recognise a mine as a foe rather than a friend, and press the IFF button in order to change the status of the weapon system (RTIFF). Thus, rather than an effect on memory load, this may reflect the efficiency with which subjects could keep track of several secondary aspects of the game and act on the information available. An effect on the efficient use of the bonus system (BONADVAN) would support this view. However, the effect of word span on BONADVAN was very small and may have been due to chance ($CR = 4.411$; $p < 0.1$).

Therefore, it appears that two major aspects of game performance were identified by secondary task methodology, namely accuracy and timing of responses, and the more complex monitoring of several sources of information. However, there were a number of component measures that were somewhat affected overall by secondary tasks, but not by any individual task. There were three measures involved: the number of times the ship moved into the line of fortress fire (NMOVIN: $F(8, 100) = 2.674$; $p < 0.05$), the total amount of movement around the screen (MOVMNT: $F(8, 100) = 3.568$; $p < 0.01$), and the number of times the subject attempted an inappropriate double shot at the fortress (NBDBLE: $F(8, 100) = 3.452$; $p < 0.01$). The first two of these measures should reflect perceptual-motor control which is likely to be an important aspect of game performance. However, largely due to the intensive testing schedule for experiment 1, none of the secondary tasks were chosen specifically to study control of ship movement, and a second experiment was carried out to investigate this issue.

Finally, there were two potential sources of artifact in the procedure used for two of the secondary tasks: rapid tapping and repeat-a-day. The rapid tapping task appeared to produce substantial impairments in game performance. However, it is possible that these arose from physical disruption of the lap-held control board by the continual leg movement that was required in this condition. This would be true to a lesser extent of slow tapping and tapping-to-a-day. In order to examine this possibility, the rapid tapping condition was repeated in experiment 2, and the control board was placed on a table between the subject and the display. In addition, in the previous experiment there was a tendency for the foot pedals to move away from the subject while tapping, despite a set of friction pads designed to prevent this. Therefore, the foot pedals were firmly fixed to the floor for the second study.

A possible second source of artifact lay in the repeat-a-day and next day conditions. The days of the week were presented by a Votrax voice synthesizer through a loudspeaker. The days of the week could be identified fairly easily, however, the speech quality was poor. Therefore, the slowing of response times to this stimulus in the dual task conditions, may have been due to a general lack of intelligibility, which was emphasised when combined with the game (including sound effects). As such, the repeat-a-day task was also included in experiment 2. However, the days of the week were presented as digitized speech through headphones, yielding substantially better quality speech than that produced by the Votrax voice synthesizer.

Experiment 2

Method

The set of secondary tasks chosen for experiment 2 were selected largely to study the relation between perceptual-motor control in the game and visuo-spatial processing functions. A typical experiment within the working memory tradition, involves two contrasting tasks, one of which has a substantial visuo-spatial processing component, while the other involves primarily verbal processing. Two sets of tasks were chosen for present purposes. While the verbal/visuo-spatial contrast was maintained in each set, the major difference between the sets was in terms of the relative general processing load required by each. These are described below.

Secondary tasks

(1) Brooks Visuo-spatial task. This was based on a task originally reported by Brooks (1967), that involved the subject retaining a sequence of movements through a 4 × 4 matrix pattern. Subjects were given a starting square which was always the square in the second row and second column. They were then given a sequence of eight sentences describing a path through the matrix, with consecutive numbers in adjacent squares. An example might be:

In the starting square put a 1.
In the next square to the right put a 2.
In the next square to the right put a 3.
In the next square down put a 4.
In the next square down put a 5.
In the next square to the left put a 6.
In the next square up put a 7.
In the next square to the left put an 8.

At the end of this sequence, the subjects were required to repeat the sequence verbatim, and to use a visual image of the path through the matrix as the means for retention. The sentences were presented as digitized speech (at 10KHz), though headphones, and accuracy of recall was recorded by the experimenter. When combined with the game, subjects were presented with the secondary task continuously. Presentation rate was one sentence every three seconds, with twenty seconds allowed for recall, before the start of the next sequence. Performance on this task is typically disrupted by visual tracking tasks such as a pursuit rotor, suggesting an overlap between perceptual-motor control and the visuo-spatial content of the main task (e.g. Baddeley et al. 1975b). Tracking tasks interfere much less, or not at all with a verbal version of this task, which is described below.

(2) Brooks verbal task. In this version of the task, the subject again hears a sequence of sentences, but with the words 'good', 'bad', 'slow' and 'quick' substituted for directions in the visuo-spatial version. It was emphasised that there was no intended

relationship between the two sets of words, and subjects were encouraged to use verbal rehearsal in order to retain the sequence. Performance on this version of the task is generally poorer, and in common with most other studies, subjects were given six sentences rather than eight in an attempt to equate performance between the conditions. Sentences were presented at a rate of one every three seconds, and 15 seconds were allowed for recall.

(3) Map task. This involved presenting subjects with a map of an island, on which were marked six locations. On a given map, the names of the locations were chosen from a particular category. For example, one map contained the names of well-known scientists such as 'Darwin Gardens' or 'Newton Monastery'. The name and type of location (e.g. Monastery, Airport etc.) were unique to a particular map and location, and no location was directly north, south, east or west of any other location. Subjects were given one minute to study the map and were then asked questions concerned with the relative direction between pairs of locations, for example:

Is Newton Monastery north of Darwin Gardens?
Is Einstein Hospital southwest of Curie Stables?
Is Bell Lake east of Newton Monastery?

In the first example shown, Newton Monastery could be northeast or northwest of Darwin Gardens for the correct response to be positive. Questions were presented through headphones as digitized speech, responses were timed by a voice key, from the onset of the first word of the question, and error responses were recorded by the experimenter. When combined with the game, the subjects were given the map prior to the start, and then presented with the questions continuously. Each question took approximately two seconds to present, and subjects were given three seconds in which to respond, before the next question.

(4) Limerick task. This was designed as a verbal equivalent of the map task, and involved the subject learning a short poem in the form of a limerick. One example was:

There was a young thief of repute,
Whose ideas were surprisingly cute.
He built a balloon,
By the light of the moon,
And used it to gather his loot.

The subject was then given a number of questions concerned with the relative position of various words in the poem, for example:

Does thief come before surprisingly?
Does light come after gather?
Does built come after ideas?

As with the map, subjects were given one minute to learn the limerick and the questions were presented as digitized speech through headphones. Responses were timed with a voice key, and error responses were recorded by the experimenter.

The remaining secondary tasks for this study were chosen to ensure that some of the results of experiment 1 were not due to artifacts of the procedures adopted.

(5) Rapid tapping. The procedure for this task was identical to that for experiment 1, except that the control board was placed on a table between the subject and the display and the foot pedals were firmly fixed to the floor.

(6) Repeat-a-day. The procedure for this task was also virtually identical to that used in experiment 1, except that the days of the week were presented as digitized speech (at 10KHz) through headphones. The game sound effects were also presented through headphones, with the sound balance adjusted to ensure that the effects were at a volume sufficiently low to avoid disrupting the intelligibility of the speech.

General procedure

Subjects were tested over three sessions, with a gap of at least one day between sessions. As with experiment 1, the first session was used for practice and control data. Subjects were given short practice sessions, followed by five minutes on each of the first four secondary tasks, and two minutes on rapid tapping and repeat-a-day. The conditions were given in the order: game alone, map, limerick, repeat-a-day, Brooks spatial, Brooks verbal, Rapid tapping and game alone once more.

In the subsequent two sessions, subjects were given five minutes on each secondary task alone for map, limerick, Brooks spatial and verbal, and two minutes on repeat day and tapping, followed in each case by combination with a full game. As subjects were already well practiced on the game alone, neither of these sessions included this condition. In addition, the testing sessions were already fairly lengthy and intensive. The order of conditions for session two was otherwise identical to that for session one, while this order was reversed for session three. As before, subjects were asked to rate the difficulty of each condition on a ten-point scale, and were asked to recall the foe letters at the end of each game. Subjects were given a short rest half way through each session, each of which lasted about two hours.

Subjects

Five of the subjects who took part in experiment 1 also took part in experiment 2. The sixth subject was not available and was replaced by another subject whose age, level of training and typical score were equivalent to those in the rest of the group. This experiment took place approximately three months after experiment 1.

Results

As with experiment 1, effects on mean game score and secondary task performance are reported first, followed by an examination of effects on game component measures.

The overall analysis of variance on game score suggested there was a significant effect of our secondary tasks ($F(6, 30) = 6.91$; $MSe = 1139012$; $p < 0.01$). Summary data for repeat-a-day and rapid tapping are shown in table 7. These results largely replicate those found in experiment 1, with no effect of repeat-a-day on game score, but with longer response times to the secondary stimulus ($t = 2.82$, $df = 28$; $p < 0.01$), and a deterioration in game score with rapid tapping ($CR = 7.95$; $p < 0.01$), but with no

Table 7
Mean game score and response time (msec) for game alone, repeat a day and rapid tapping in experiment 2.

	Control	Repeat day	Rapid tapping
Game score	5972	5592	4770 [a]
Sec. task alone (msec)		657	580
Sec. task + game (msec)		928	644
Difference (msec)		270 [a]	64
Mean difference in *SD* (msec)		152 [a]	77

[a] $p < 0.01$.

effect on inter-tap intervals. Only response times for repeat-a-day proved to be significantly more variable when combined with Space Fortress ($t = 4.41$, $df = 28$; $p < 0.01$).

Mean scores for the map, limerick and Brooks tasks are shown in table 8. Game score was significantly affected by all of these secondary tasks; ($CR = 9.839$; $p < 0.01$) for the map task, ($CR = 7.189$; $p < 0.01$) for the limerick task, ($CR = 13.996$; $p < 0.01$) for Brooks spatial and ($CR = 12.792$; $p < 0.01$) for Brooks verbal task. Response times to the map and limerick questions were not significantly affected by dual task, and neither measure was more variable. Neither was there an effect of dual task on the

Table 8
Mean game score and response time (msec) for game alone, map, limerick and Brooks spatial and verbal tasks.

	Control	Map	Limerick	Brooks spatial	Brooks verbal
Game score	5972	4474 [a]	4797 [a]	3809 [a]	3998 [a]
Sec. task alone (msec)		4199	3982	7.43	5.23
Sec. task + game (msec)		4310	4223	5.83	4.19
Difference (msec)		111	241	2.60 [a]	1.04 [a]
Mean difference in *SD* (msec)		110	178	–	–

[a] $p < 0.01$.

Table 9
Percentage decrements in game score and secondary task performance, and mean rated difficulty for all conditions in experiment 2.

	Decrement in game score (%)	Rated difficulty game + task [c]	Decrement in sec. task (%)	Rated difficulty task alone [d]
Game alone	–	2.42		
Repeat day	6.36	3.58 [a]	41.10 [b]	1.22 [b]
Limerick	19.68 [b]	5.50 [b]	6.05	3.78 [b]
Rapid tap	20.13 [b]	4.33 [b]	11.03	1.72 [b]
Map	25.08 [b]	6.25 [b]	2.64	4.61 [b]
Brooks verbal	33.05 [b]	7.83 [b]	19.89 [b]	5.67 [b]
Brooks spatial	36.22 [b]	7.42 [b]	34.99 [b]	4.72 [b]

[a] $p < 0.05$; [b] $p < 0.01$.
[c] Significance levels are for comparison with rating for game alone.
[d] Significance levels are for comparison with rating for secondary task with the game.

number of errors or omitted responses in these conditions. Performing Space Fortress significantly reduced the mean number of items recalled for each of the Brooks tasks ($t = 4.61$, $df = 11$; $p < 0.01$) for spatial material and ($t = 5.59$, $df = 11$; $p < 0.01$) for the verbal material.

Although there was a tendency for the map task to produce larger decrements than the limerick task, and for Brooks spatial to produce larger decrements than Brooks verbal, these differences were not statistically reliable.

Table 9 summarises the effects on game score and secondary task performance in terms of percentage decrements, along with mean rated difficulty. Unlike experiment 1, rated difficulty on the secondary tasks when they were performed alone appeared to be associated with the largest decrements in game score ($r = 0.861$; $p < 0.05$), and rated difficulty of the dual tasks ($r = 0.968$; $p < 0.01$). The one anomaly was rapid tapping, which appeared to produce a relatively large impairment in game performance (20.13%) but received a very low difficulty rating when performed alone.

There were two general aims of this experiment: first to investigate the possibility of artifacts in some of the results of experiment 1, and second, to investigate the role of visuo-spatial processing in the game. The results on rapid tapping and repeat-a-day were replicated in this study, supporting the view that generating a secondary mechanical response appears to require resources that overlap with some crucial aspects of game performance. In addition, adequate performance on Space Fortress appears to be associated with slower responses to an external stimulus. The results of experiment 2 suggest that neither of these results could be accounted for in terms of physical disruption of the game controls by tapping, or lack of intelligibility of the auditory stimulus.

The results of the remaining secondary tasks, provide a less clear picture. Although, all of the tasks were associated with impairments in performance, there was no clear differential disruption by the spatial as opposed to the verbal tasks. The relatively larger disruption by the Brooks tasks can easily be accounted for in terms of general processing load, since both of these tasks were rated as the most difficult of the set. However, we have already shown that the ratings of difficulty tend not to be reliable predictors of task performance. It is possible that this lack of differential impairment results from combining game component measures in a single score, and this is explored below.

Effects on game components

The scores for each game component in each experimental condition were entered into separate analyses of variance. Table 10 shows the mean scores for each game component for game alone and with each of the secondary tasks. Table 11 shows a summary of the results of these analyses in terms of the effects of particular secondary tasks on groups of component measures. As with experiment 1, we adopted a more conservative criterion of statistical reliability and only those effects that were significant at the 1% level or better were included. Also as before, we will report only the post hoc analyses using Newman-Keuls tests for those components where the overall analysis of variance was significant at the 1% level or better.

From table 11, it appears that as with experiment 1, repeat-a-day affected none of the game components. Again as before, rapid tapping affected number of fortress hits (NFOHITS: $CR = 6.323$; $p < 0.01$) and number of fortress destructions (NFODES: $CR = 5.42$; $p < 0.01$). This task also affected shooting efficiency (SHTEFF: $CR = 6.44$; $p < 0.01$). The limerick task affected the response time to the presence of a foe mine (RTIFF: $CR = 6.439$; $p < 0.01$), and the time to kill the fortress, or a foe mine

Table 10
Mean scores on component measures for game alone and with each secondary task in experiment 2.

Game component	Control	Repeat-a-day	Rapid tap	Brooks verbal	Brooks spatial	Limerick	Map
MOVMNT	732	856	893	876	862	856	862
NBDBLE	7.08	6.58	2.83	3.58	2.83	4.08	4.00
NBDIFF	0.10	0.10	0.36	0.52	0.10	0.47	0.67
NBDINT	0.25	1.00	1.58	2.42	1.42	0.75	1.25
NFODES	36.08	34.92	28.92	25.75	26.17	30.50	29.17
NFOHITS	432	420	335	318	321	371	352
NMOVIN	1.92	5.42	9.83	4.75	4.33	6.08	4.08
NSHDMG	2.33	4.67	6.42	8.58	10.42	7.08	6.83
NWRAP	0.25	0.50	0.58	0.67	1.08	0.08	0.50
PCNTBON	95.8	94.2	85.0	63.3	55.8	80.0	65.0
RTIFF	85	98	93	114	122	115	118
SHTEFF	89.8	89.7	84.0	89.0	87.3	88.4	87.6
TIMKLFO	183	198	202	214	221	210	211
TIMKLFR	160	165	165	173	182	169	168
TOTSHOTS	531	514	447	407	414	468	453

Table 11
Effects of secondary tasks on component measures in experiment 2.

Secondary task	Components affected
Repeat-a-day	None
Rapid tapping	NFOHITS, NFODES, SHTEFF, TOTSHOTS
Limerick	RTIFF, TIMKLFO, TOTSHOTS
Map	RTIFF, TOTSHOTS
Brooks verbal	NBDINT, NFOHITS, NFODES, RTIFF, TIMKLFO, TOTSHOTS
Brooks spatial	NFOHITS, NFODES, RTIFF, TMKLFO, NSHDMG, TOTSHOTS

(TIMKLFO: $CR = 5.184$; $p < 0.01$). The map task affected only RTIFF ($CR = 7.012$; $p < 0.01$). The Brooks verbal task affected the number of bad intervals in a double press of the IFF button (NBDINT: $CR = 7.239$; $p < 0.01$), the number of hits on a foe (NFOHITS: $CR = 7.42$; $p < 0.01$), the number of fortress destructions (NFODES: $CR = 7.815$; $p < 0.01$), RTIFF ($CR = 6.189$; $p < 0.01$) and TIMKLFO ($CR = 6.001$; $p < 0.01$). The Brooks spatial task affected NFOHITS ($CR = 7.21$; $p < 0.01$), NFODES ($CR = 7.50$; $p < 0.01$), RTIFF ($CR = 7.906$; $p < 0.01$), TIMKLFO ($CR = 7.243$; $p < 0.01$) and NSHDMG ($CR = 6.416$; $p < 0.01$).

Discussion

One fairly striking observation from these data is that none of the measures of tracking control (NWRAP, NMOVIN, MOVMNT) were affected by any of our secondary tasks, in particular, the visuo-spatial memory tasks. The Brooks spatial task was affected by simultaneous game performance, but so too were other secondary tasks. In addition, the alternative visuo-spatial task (maps) was largely unaffected by dual task. Indeed, what is most striking about both of the Brooks tasks is the range of game components which they affect. This supports the view that the interference produced by these tasks results from general rather than specific processing.

However, this begs the question as to the processes that may underlie game performance. The specific disruption of visuo-spatial processing, including tracking a moving target, by the Brooks task is well established (Brooks 1968; Baddeley and Lieberman 1980). This suggests that the cognitive processes involved in tracking on a pursuit rotor or following a randomly moving target (e.g. Baddeley et al. 1986) are not involved in control of ship movement in Space Fortress. This seems incongruous as the tasks superficially seem very similar.

There are two possible explanations for this in terms of differences between these and earlier studies. First, the subject is controlling the movement of a shape on a VDU screen. He is not attempting to track a moving target over which he has no control. The control aspects of the task would most likely give rise to the general interference effects found in experiments 1 and 2. Within the working memory framework, this would involve the central executive component. However, there is still a substantial 'tracking' component to the task. For example, the subject has to track the movement of mines on the screen in order to fire missiles at them. We would therefore still expect to get some differential disruption by a concurrent visuo-spatial task, in addition to any general load effect.

A more plausible explanation lies in the level of expertise involved. In more traditional studies, the subjects are given sufficient practice to be familiar with the task. This is generally of the order of several minutes. The task in Space Fortress is somewhat more complex, however, all of our subjects had a minimum of 20 hours practice on the game prior to the introduction of our secondary tasks, and the amount of practice may well be crucial. One possibility is that the visuo-spatial tracking elements of the task are largely automatic (e.g. Jonides et al. 1985; Shiffrin and Schneider 1977) in experts on the game. In experiment 3 we attempted to test this possibility by studying the effects of our secondary tasks on the performance of subjects who had relatively little practice on Space Fortress.

Experiment 3

Method and procedure

The aim in experiment 3 was to investigate the effects of our various secondary tasks on performance during the early stages of training. We chose those of our tasks which appeared to yield the most interesting results with experts, namely the tasks used in experiment 2. In particular we were interested in whether visuo-spatial or verbal memory loads would have differential effects on game performance in novices, where this had not occurred in experts. The basic procedure for each task and its combination with the game was more or less identical to that adopted for experiment 2.

The general procedure was designed to examine effects on performance at two different stages in training. Subjects were first given three hours of practice on the game. Next, they were given one session consisting of practice on each of the secondary tasks. There then followed two sessions where each secondary task was performed alone and with the game. Next, subjects were given a further five hours practice on the game alone, followed by a further two sessions involving the secondary tasks.

Subjects were nine students at the University of Illinois at Urbana-Champaign. None had played Space Fortress prior to this experiment.

Results

Mean game scores are shown in table 12 for the game alone and when played with each of the secondary tasks. The control game score was based on the mean of four games at each stage of training. Two of the games were played immediately prior to the introduction of secondary tasks and two were played immediately after the secondary tasks were completed. These data are shown separately for after three hours of training and after eight hours of training.

Table 12
Mean game score, response times and recall scores for game alone and for each of the secondary tasks used in experiment 3.

	Control	Repeat day	Rapid tap	Map	Limerick	Brooks spatial	Brooks verbal
After 3 hours practice							
Game score	1244	1200	987	315 [a]	711 [a]	360 [a]	289 [a]
RT and recall score (alone)		566	557	4268	3992	7.31	4.76
RT and recall score (with game)		930	534	4772	4303	4.23	3.29
Errors (alone)		0.06	–	3.22	3.00	–	–
Errors (with game)		2.00	–	9.83	5.94	–	–
After 8 hours practice							
Game score	2938	2713	2146 [a]	1940 [a]	2192 [a]	1459 [a]	1559 [a]
RT and recall score (alone)		538	547	4396	3913	7.43	5.12
RT and recall score (with game)		858	521	4593	4215	5.04	3.49
Errors (alone)		0	–	3.72	1.89	–	–
Errors (with game)		0.67	–	8.00	5.50	–	–

Note: Secondary task data for different levels of training were analysed in a single ANOVA. See text for significance levels.
[a] $p < 0.01$.

The game scores for each subject were entered into an analysis of variance with three factors treated as repeated measures, namely three versus eight hours of training (2) × first versus second replication of secondary tasks (2) × game alone versus game plus each of six secondary tasks (7). The analysis revealed that level of training had an overall effect on game score ($F(1, 8) = 71.4$; $MSe = 1744809$; $p < 0.001$), and that game score varied according to secondary tasks ($F(6, 48) = 23.2$; $MSe = 349211$; $p < 0.001$). The effect of the secondary tasks was contingent upon the level of training as shown by an interaction between these variables ($F(6, 48) = 2.32$; $MSe = 208169$; $p < 0.05$).

There was an effect of replicating the secondary tasks ($F(1, 8) = 11.8$; $MSe = 267972$; $p < 0.01$). However, the effect of replication only occurred at the early stage of training, where the mean game score for block one was 522 and for block two was 936. After eight hours of training, the block one game score was 2118 and the block two game score was 2152. This interaction was significant ($F(1, 8) = 17.21$; $MSe = 132573$; $p < 0.01$). There was also a tendency for the effect of conditions to change with replications ($F(6, 48) = 3.40$; $MSe = 206548$; $p < 0.01$).

Multiple pairwise comparisons on the game scores with each secondary task were carried out using Newman-Keuls tests. A summary of the significance levels of these comparisons is shown in table 12. Early in training, when compared with playing the game alone, game score was significantly disrupted by the limerick task ($CR = 4.77$), the map task ($CR = 8.07$), the Brooks verbal task ($CR = 8.29$) and the Brooks spatial task ($CR = 7.70$). It was not affected by repeating a day of the week or foot tapping. Later in training, game score was still disrupted by the limerick task ($CR = 4.91$), the Map task ($CR = 6.83$), the Brooks verbal task ($CR = 9.73$) and the Brooks spatial task ($CR = 10.49$). However, performance at this level of training was also significantly disrupted by foot tapping ($CR = 5.26$).

We also examined the data for the two replications of the secondary task procedure at the earlier stage of training. An analysis of variance supported the initial analysis in that there was a significant effect of secondary tasks overall ($F(6, 48) = 13.04$; $MSe = 242705$; $p < 0.001$). There was a significant increase in mean overall game score from block one to block two ($F(1, 8) = 60.92$; $MSe = 88713$; $p < 0.001$), and a significant change in the pattern of secondary task disruption from block one to block two. A similar analysis on the data from a later stage of training showed only an effect of secondary tasks ($F(6, 48) = 17.24$; $MSe = 314674$; $p < 0.001$). However, repeating the procedure had no effect on game score ($F < 1$), nor did these variables interact ($F < 1$).

Multiple comparisons at the earlier stage of training showed that during the first block with secondary tasks, game score was significantly affected by the limerick task ($CR = 4.48$, the map task ($CR = 8.44$), Brooks verbal task ($CR = 6.38$) and Brooks spatial task ($CR = 5.11$). In addition, the map task resulted in significantly poorer game performance than did the limerick task ($CR = 3.96$).

For the second block of secondary tasks, the pattern was slightly different, with game score affected by the map task ($CR = 4.0$) and the Brooks verbal ($CR = 5.88$) and Brooks Spatial task ($CR = 6.04$), but not by the limerick task.

We did not carry out a similar analysis at the later stage in training as the pattern of disruption associated with our secondary tasks did not differ across replications.

We also carried out separate analyses of variance on data from each of the secondary tasks. Mean response times for repeat-a-day, rapid tapping, map and limerick questions are shown in table 12, along with recall scores for the Brooks spatial and verbal tasks. Error data for repeat-a-day, map and limerick tasks are shown in the same table.

An analysis of variance on correct response times showed that playing the game significantly slowed response time to days of the week ($F(1, 8) = 66.97$; $MSe = 31576$; $p < 0.001$), and response times decreased with training ($F(1, 8) = 23.95$; $MSe = 1881$; $p < 0.005$). A similar analysis on the error data, suggested that playing the game produced a small increase in the number of errors in this task ($F(1, 8) = 6.78$; $MSe = 4.524$; $p < 0.05$). Training reduced the number of errors overall ($F(1, 8) = 7.16$; $MSe = 1.212$; $p < 0.05$), and the effect on errors of playing the game was reduced with training ($F(1, 8) = 5.58$; $MSe = 1.316$; $p < 0.05$). However, it is clear from table 12, that the absolute numbers of errors involved was very small.

Response times (inter-tap intervals) in the foot tapping tasks were unaffected by game playing ($F < 1$) or by training ($F(1, 8) = 1.58$).

Response times to questions about limericks significantly increased when combined with the game, ($F(1, 8) = 63.85$; $MSe = 51770$; $p < 0.001$), as did the number of error responses ($F(1, 8) = 16.35$; $MSe = 11.826$; $p < 0.01$). However, training affected neither measure, nor did it affect the effect on performance of playing the game.

There was a marginal effect on response time for map questions with the game ($F(1, 8) = 4.70$; $MSe = 566612$; $p < 0.07$), and a suggestion that this interacted with training ($F(1, 8) = 4.80$; $MSe = 132726$; $p < 0.06$). There was a more substantial effect on number of errors to the map questions ($F(1, 8) = 14.79$; $MSe = 36.087$; $p < 0.01$), and replication of the map task reduced the number of errors observed ($F(1, 8) = 8.49$; $MSe = 5.503$; $p < 0.05$). However, training had no effect on performance with this task.

The Brooks spatial task was disrupted by game playing, as shown by a drop in the number of directions correctly recalled ($F(1, 8) = 85.10$; $MSe = 1.671$; $p < 0.001$). Training had no effect overall, but there was a suggestion that the effect of game playing was marginally smaller with more training ($F(1, 8) = 4.22$; $MSe = 0.328$; $p < 0.08$). The Brooks verbal task was unaffected by game playing ($F(1, 8) = 1.18$) or by training, ($F < 1$).

We next considered the subjective ratings of difficulty. Mean difficulty ratings along with percentage changes in game and secondary task performance are shown in tables 13 and 14 for early in training and later in training respectively. Three way analyses of variance (early or later in training, trial block 1 or 2 and secondary task alone or with the game) were carried out on difficulty ratings for each secondary task in turn.

Difficulty ratings for repeat-a-day increased when this task was combined with the game ($F(1, 8) = 56.82$; $MSe = 1.097$; $p < 0.001$), but the task was not viewed as more or less difficult as a result of trial blocks or training. Rapid tapping was also viewed as more difficult with the game ($F(1, 8) = 62.23$; $MSe = 2.750$; $p < 0.001$). However, the task itself was thought to be less difficult with training ($F(1, 8) = 3.36$; $MSe = 1.597$; $p < 0.05$), and as a result of trial blocks ($F(1, 8) = 6.07$; $MSe = 0.826$; $p < 0.05$). The effect of the game on perceived difficulty was reduced with training ($F(1, 8) = 10$;

Table 13
Percentage decrements in game score and secondary task performance, and mean rated difficulty for all conditions after three hours of training in experiment 3.

	Decrement in game score (%)	Rated difficulty game + task	Decrement in sec. task (%)	Rated difficulty task alone
Game alone	–	1.89		
Repeat day	3.54	3.333	6.43	1.333
Rapid tap	20.66	5.833	4.1	2.889
Limerick	42.85	6.167	7.79	3.944
Brooks spatial	71.06	8.500	42.1	4.333
Map	74.68	6.722	11.8	4.278
Brooks verbal	76.77	8.611	30.9	5.500

Note: Difficulty ratings for different levels of training were analysed in a single ANOVA. See text for significance levels.

Table 14
Percentage decrements in game score and secondary task performance, and mean rated difficulty for all conditions after eight hours of training in experiment 3.

	Decrement in game score (%)	Rated difficulty game + task	Decrement in sec. task (%)	Rated difficulty task alone
Game alone	–	1.89		
Repeat day	7.66	3.000	59.5	1.278
Limerick	25.39	4.889	7.7	3.000
Rapid tap	26.96	5.111	4.8	1.889
Map	33.97	5.944	4.5	3.833
Brooks verbal	46.94	7.389	31.8	4.500
Brooks spatial	50.34	6.778	32.2	3.389

Note: Secondary task data for different levels of training were analysed in a single ANOVA. See text for significance levels.

$MSe = 0.312$; $p < 0.05$). This is an interesting contrast to the performance data, which indicated that the game was more disrupted by rapid tapping after extra training.

Playing the game with the limerick task was seen as more difficult than the limerick task alone ($F(1, 8) = 53.17$; $MSe = 1.431$; $p < 0.001$), and the overall difficulty decreased with training ($F(1, 8) = 15.09$; $MSe = 1.472$; $p < 0.01$), but was not affected by trial blocks. However, training interacted with dual versus single task performance ($F(1, 8) = 6.40$; $MSE = 0.868$; $p < 0.05$) such that after further training the dual task performance was not seen as any more difficult than limerick task alone.

The map task was rated as less difficult than game and map task together ($F(1, 8) = 24.41$; $MSe = 3.826$; $p < 0.01$). Training had no effect on difficulty, and training did not interact with the effect of the dual-task condition.

Performing the game with Brooks spatial task was rated as more difficult than the Brooks task alone ($F(1, 8) = 127.56$; $MSe = 2.014$; $p < 0.001$). Overall difficulty was rated as lower after further training ($F(1, 8) = 26.95$; $MSe = 1.188$; $p < 0.001$), but these variables did not interact. Finally, Brooks verbal task, combined with the game was rated as more difficult than the verbal task alone ($F(1, 8) = 78.55$; $MSe = 2.06$; $p < 0.001$), but training had only a marginal effect overall ($F(1, 8) = 4.77$; $MSe = 4.66$; $p < 0.07$), and the variables did not interact. Trial blocks showed a small decrease in overall difficulty ($F(1, 8) = 7.53$; $MSe = 1.063$; $p < 0.05$).

Rated difficulty for each of the secondary tasks when performed without the game, appeared to be associated with decrements in game score in the dual task conditions at an early stage of training ($r = 0.933$, $p < 0.01$), but not with decrements in the secondary tasks ($r = -0.395$). At a later stage of training the picture was less clear. The correlation between rated difficulty of the task alone and decrements in game score was significant ($r = 0.833$, $p < 0.05$), but as with experiments 1 and 2 rapid tapping proved to be an exception, being rated as a very easy task on its own, yet producing a significant decrement in game performance. For example, it was rated as easier than the limerick task when performed alone but produced about the same decrement in game score.

Effects on game components

The sets of scores for each game component in each experimental condition were entered into separate analyses of variance. Table 15 shows mean values for each component measure for the game alone and for the game with each secondary task as measured early in training. Table 16 shows the same class of data for the later stage of training. Table 17 shows a summary of the results of the analysis in terms of the effects of particular secondary tasks on groups of component measures. As with experiments 1 and 2, a conservative criterion of 1% was set for tests of statistical reliability. Also as with the earlier experiments, we were primarily interested in the effects of specific secondary tasks on each component and will report only post hoc analyses on those components for which the overall analysis of variance was significant at the 1% level or better.

As with the earlier experiments, repeat-a-day had no effect on any of the game components that we examined. Unlike the results for experts, at an early stage of training tapping reduced the total number of shots fired ($CR = 5.269$; $p < 0.01$), but affected no other variables. Late in training it affected a wider range of components, reducing the number of hits on the fortress (NFOHITS; $CR = 9.359$), the number of times the fortress was destroyed (NFODES: $CR = 9.056$), and shooting efficiency

Table 15
Mean scores on component measures for game alone and with each secondary task after 3 hours of training in experiment 3.

Game component	Control	Repeat-a-day	Rapid tap	Brooks verbal	Brooks spatial	Limerick	Map
MOVMNT	1636	1646	1641	1667	1646	1651	1666
NBDBLE	0.917	1.017	0.722	1.167	0.778	0.944	0.778
NBDIFF	0.444	0.833	0.333	1.333	0.556	0.667	1.667
NBDINT	3.556	4.389	5.778	3.778	3.222	3.611	2.722
NFODES	5.694	5.556	4.278	2.889	3.222	4.056	3.500
NFOHITS	79.69	85.67	62.89	55.33	55.00	64.22	58.56
NMOVIN	39.50	42.11	43.67	42.11	40.72	39.44	41.56
NSHDMG	6.86	7.17	7.06	9.78	9.28	9.17	11.00
NWRAP	36.14	32.72	42.39	37.89	43.06	38.72	42.83
PCNTBON	73.61	73.33	66.67	42.22	43.89	60.00	47.78
RTIFF	163	146	159	207	196	190	242
SHTEFF	59.35	60.23	56.37	53.04	54.39	57.71	54.39
TIMKLFO	392	377	410	429	417	402	434
TIMKLFR	240	236	249	262	264	266	292
TOTSHOTS	186	196	166	153	147	162	153

Note: Data for different levels of training were analysed in a single ANOVA. See text for significance levels.

Table 16
Mean scores on component measures for game alone and with each secondary task after 8 hours of training in experiment 3.

Game component	Control	Repeat-a-day	Rapid tap	Brooks verbal	Brooks spatial	Limerick	Map
MOVMNT	1656	1664	1659	1699	1684	1676	1658
NBDBLE	1.944	2.667	0.833	0.667	0.500	1.167	1.056
NBDIFF	0.833	0.500	0.667	1.056	0.500	0.889	0.778
NBDINT	3.056	2.722	4.167	2.222	3.944	3.167	2.667
NFODES	13.47	12.50	8.94	7.17	7.11	10.06	9.44
NFOHITS	168	162	110	99.6	96.8	131	122
NMOVIN	31.17	34.56	33.78	31.83	29.39	31.56	31.11
NSHDMG	5.19	5.56	6.00	7.61	7.11	6.56	6.89
NWRAP	16.03	15.61	22.89	19.33	19.33	16.11	17.56
PCNTBON	87.22	90.00	85.56	58.89	47.78	79.44	68.33
RTIFF	116	114	126	147	155	142	159
SHTEFF	77.68	76.84	71.47	74.48	72.89	77.13	75.95
TIMKLFO	290	285	309	323	345	295	327
TIMKLFR	205	197	205	248	230	216	222
TOTSHOTS	261	261	202	176	177	213	203

Note: Data for different levels of training were analysed in a single ANOVA. See text for significance levels.

(SHTEFF; $CR = 7.086$). With experts all of these components were also disrupted by rapid tapping. In addition, total shots were reduced at this level of training (TOTSHOTS; $CR = 6.669$).

Early in training the limerick task disrupted only the percentage of times the bonus option was taken (PCNTBON; $CR = 4.298$). Later in training, this task affected rather more components, specifically reducing the number of hits on the fortress (NFOHITS; $CR = 5.971$), the number of fortress destructions (NFODES; $CR = 6.832$), and total shots (TOTSHOTS; $CR = 5.45$). It did not affect PCNTBON at this level of training.

The map task produced a rather different pattern of disruption. Early in training it increased the number of times the ship was damaged by a mine (NSHDMG; $CR = 6.479$), reduced the number of times that the fortress was destroyed (NFODES; $CR = 5.617$), delayed the response time to pressing the IFF button when a foe mine appeared (RTIFF; $CR = 7.229$), and reduced the percentage of occasions that the bonus option was taken (PCNTBON; $CR = 8.158$). Later in training it reduced fortress hits (NFOHITS; $CR = 7.443$), fortress destructions (NFODES; $CR = 8.056$) (RTIFF; $CR = 6.903$), bonus selection (PCNTBON; $CR = 4.894$) and total number of shots fired (TOTSHOTS; $CR = 9.65$). In addition it lengthened the mean interpress interval (MEANINT; $CR = 5.791$), and response time to a foe mine (RTIFF; $CR = 6.903$). With experts, only the response time to a foe mine was disrupted by this task.

Early in training the Brooks verbal task reduced the number of fortress hits (NFOHITS; $CR = 5.232$), the number of times the fortress was destroyed (NFODES;

Table 17
Effects of secondary tasks on component measures in experiment 3.

Secondary task	Components affected early in training	Components affected later in training
Repeat-a-day	None	
Rapid tapping	TOTSHOTS	TOTSHOTS, NFOHITS, NFODES SHTEFF
Limerick	PCNTBON	NFOHITS, NFODES, TOTSHOTS
Map	NSHDMG, NFODES, RTIFF, PCNTBON, MEANINT	NFODES, RTIFF, PCNTBON, MEANINT, NFOHITS, TOTSHOTS
Brooks verbal	NFOHITS, NFODES, PCNTBON, TOTSHOTS	NFOHITS, NFODES, PCNTBON TOTSHOTS, RTIFF
Brooks spatial	NFOHITS, NFODES, KLTIMFR, PCNTBON, TOTSHOTS	NFOHITS, NFODES, KLTIMFR, PCNTBON, TOTSHOTS, RTIFF KLTIMFO, RESMANG, MEANINT

$CR = 6.376$), the selection of the bonus option (PCNTBON; $CR = 9.913$), and total number of shots fired (TOTSHOTS; $CR = 6.050$). Later in training it reduced values on these same variables (NFOHITS; $CR = 10.965$), (NFODES; $CR = 12.610$), (PCNTBON; $CR = 7.340$), (TOTSHOTS; $CR = 9.701$). In addition, the response time to a foe mine was increased (RTIFF; $CR = 4.951$). Experts were disrupted on NFOHITS, NFODES and RTIFF by this task.

For the Brooks spatial task, early in training there was a reduction in the number of fortress hits (NFOHITS; $CR = 5.304$), fortress destructions (NFODES; $CR = 7.286$), selection of bonus (PCNTBON; $CR = 9.386$), as with the verbal task. However in addition, the time to energise a friendly mine was increased (KLTIMFR; $CR = 5.430$). After more training a much wider range of components was affected. There was a reduction in the number of fortress hits (NFOHITS; $CR = 11.400$), the number of fortress destructions (NFODES; $CR = 12.722$), bonus selection (PCNTBON; $CR = 10.219$), total shots (TOTSHOTS; $CR = 6.791$), and an increase in the time to energise a friendly mine (KLTIMFR; $CR = 5.507$), as was the case early in training. In addition, there was a reduction in management of available resources (RESMANG; $CR = 5.170$) and shooting efficiency (SHTEFF; $CR = 5.468$), along with an increase in time to press the IFF button after a foe mine appears (RTIFF; $CR = 6.273$), the time to destroy a foe mine (KLTIMFO; $CR = 5.019$) and the mean interpress interval on the IFF button (RTIFF; $CR = 6.273$). Experts were disrupted by this task on NFOHITS, NFODES, RTIFF and KLTIMFO.

Discussion

There are two striking, general observations from these data. The first is the differential nature of the patterns of disruption associated with different secondary tasks. At one level this is reflected in the different effects on game score as shown in table 12. Notably rapid tapping did not significantly impair game score early in training, but does so after further training. This suggests that the aspects of game performance that are impaired by generating a concurrent paced response are developed or become more important with training.

Also from table 12, the importance of visuo-spatial processing appears to show the reverse trend to that for tapping. Early in training, the map task produces significantly greater decrements in game score than does the limerick task. The difference becomes non-significant with further training. This is not true of the two Brooks tasks, where the degree of decrement associated with the verbal and the visuo-spatial tasks do not differ. However, these tasks produced a large decrement overall, and they were rated as the most difficult of the secondary tasks. This suggests that the general load effect of these tasks may have been sufficient to mask any differential effects that they may have had on game score.

The second, general observation provides more evidence on this last issue. This refers to the effect of training on the game components affected by our secondary tasks. Overall, the number of components thus affected appeared to increase with more training. This is in apparent contrast to the overall improvement in performance, as measured by game score, that was associated with greater training. From table 15, it is clear that not only did the map task result in a lower game score early in training, it also affected a wider range of game components than did the limerick task. This is also true, to a lesser extent with the Brooks tasks, where one additional factor was affected (KLTIMFR) by the spatial task.

Later in training, the Brooks tasks do appear to show differential effects. While there was no real difference in game score, clearly there were a greater number of game components affected by the visuo-spatial task.

The effects on game score of the map task suggest that visuo-spatial coding is important early in training, and this importance diminishes with further training. On the basis of the component measures, the importance of visuo-spatial coding appears to increase with training. One possible way to deal with this apparent dilemma is to consider the way in which the game score is derived.

A decrement in game score may arise for one of two reasons: either one or a few of the components that make up game score are very seriously impaired, or a range of components are affected, most or all of which feed into the total score. However, there are also several components that do not feed into game score directly. In addition, game score is determined on relatively arbitrary choices of components. This suggests that the more relevant data for this sort of comparison are those based on the component measures. It also suggests that we should consider the nature rather than the number of game components affected.

One approach would be to examine those components that were affected by particular secondary tasks, and attempt to discover whether there is anything about the components that may fit with the characteristics of the processing we believe is involved in our secondary tasks.

This does risk the danger of circularity. In addition, the component measures are likely to involve more than one underlying cognitive or physical skill, and components are unlikely to be independent of one another. In this sense, the various game components may be measures of game performance or of success of training procedures. They may be less useful as measures of the individual component skills required for that performance. Finally, there is a sense in which some of the effects of training in experiment 3 may reflect training on our secondary tasks as much as expertise in the game. At this stage therefore, our conclusions must remain tentative. However, we feel that this approach would be worthwhile to provide further insight into a rather complex set of data and to suggest avenues for further exploration.

The number of fortress destructions involves many aspects of the game, including the efficiency of dealing with mines, keeping out of the line of fire of the fortress, as well as accurate timing and aiming of shots. It is likely therefore that this measure would be sensitive to a wide range of secondary tasks. For similar reasons this should also be true of the number of hits on the fortress. Our data suggest that this is the case since at the later stage of training, both NFODES and NFOHITS are affected by all of our secondary tasks, with the one exception of repeat-a-day.

In contrast, the use of the bonus option is much more likely to reflect a general processing load, and this measure should be affected by tasks which load general working memory. From table 15, PCNTBON is affected by the limerick, map, and both Brooks tasks, all of which involve the central executive component of working memory in addition to their specifically visuo-spatial of verbal nature. PCNTBON is unaffected by rapid tapping or repeat-a-day, neither of which should involve the central executive, or general working memory capacity.

It seems reasonable to suggest that components which involve visuo-spatial tracking would be those crucial to controlling the movement of the ship. Notably, gross measures of movement were largely unaffected by the secondary tasks. These were the number of times the ship wrapped around the screen (NWRAP), the total amount of movement (MOVMNT), and the amount of movement into the line of fire of the fortress (MOVINT). More skilled aspects of tracking control are likely to arise in more sensitive measures of ship control. Thus the ability to deal with a mine would largely involve tracking the movement of the mine, then changing the status of the weapon systems if it is a foe mine, and aiming and firing a missile. Where the mine is a foe, the time to change the weapon system (RTIFF) would be a crude measure of how quickly a foe mine was recognised as such, whereas the time to destroy a mine would be a measure of tracking and aiming. Where the mine is a friend, much of the time would be in tracking and aiming, therefore we would expect that time to energise a friend mine (KLTIMFR) would be affected by a secondary task that require visuo-spatial processing. This should also be true of the time taken to destroy a foe mine, but to a lesser extent since some of the variance in this measure is due to the time taken to change the weapon system (RTIFF).

This description has some support in our data. Early in training, KLTIMFR is affected significantly by only one of our secondary tasks, namely the Brooks spatial task. Later in training, Brooks spatial task is again the only significant disrupter of this variable, and of the complementary component KLTIMFO. The time to press the IFF button in response to a foe mine is more likely to involve a general processing load, as it is thought to be handled by the central executive. RTIFF appears to involve noticing that a mine is present, remembering the letters associated with foe mines, identifying it as a foe or a friend, and then pressing the IFF button if appropriate. The attentional aspect of 'noticing' that a mine had appeared, and making a decision about what action to take are likely to be central executive type processes. RTIFF is affected by both Brooks tasks and by the map task.

The total number of shots fired is also likely to involve several aspects of game performance. It seems to be as much determined by control of the ship when firing at mines and fortress as by timing of shots at an appropriate time. It is unsurprising therefore that it is affected by all (except repeat-a-day) of our secondary tasks. Shooting efficiency should be a better measure of the extent to which shots can be aimed and fired at the appropriate time to hit a moving target. This is of course relative movement in the case of the fortress. Rapid tapping is the only factor which affects this variable (SHTEFF). This supports the view that timing and accuracy of shots is a factor that becomes crucial with increasing expertise. This would account for the large impairments found with this task in the performance of experts. An examination of tables 6 and 11 shows that SHTEFF is only affected by tapping tasks in expert performance. RTIFF is affected by the limerick, map and both Brooks tasks, as well as word span and sentence span (working memory span). Both of these variables were affected in similar ways by our group with a moderate amount of training. The same is true of NFOHITS and NFODES.

The main differences with the experts were that the time to destroy a foe mine was affected by the limerick and Brooks verbal tasks as well as the Brooks spatial task, while the time to energise a friend mine (KLTIMFR) was unaffected by any variable. This may suggest that the general load of secondary tasks, may affect the time to identify a mine as present and decide what to do about it, but that the process of tracking a mine and hitting it with a missile is relatively automatic. This in turn supports the view that control of the ship generally is relatively automatic in experts.

Overall discussion

The original aim of these studies was twofold: to provide an analysis of a complex task in terms of decomposition into component skills, and to act as a study of the use of secondary task methodology and the working memory framework in a relatively novel environment. Space Fortress provided a set of task demands which lent themselves to this type of methodology, and on the whole, the results are encouraging. The results of experiment 1 suggested that at least two different classes of skill were required; one involved in response timing and accuracy, and the other in monitoring of events and strategic control. Experiment 2 supported these conclusions. However, the lack of any differential disruption of the perceptual-motor control aspects of game performance by secondary visuo-spatial tasks was a little puzzling. A possible explanation lies in the extent of subject training. Since all of the subjects were highly practiced on the game, certain aspects of game performance may have been highly automated, and as a result much less prone to disruption.

The general aim of the Space Fortress project was to examine the learning processes during the acquisition of complex skills. Experiment 3 showed that the pattern of disruption in game performance associated with our range of secondary tasks changes as a result of training. Further, the changes were not wholly characterised by a general decrease in the disruptive effects of our secondary tasks. These results underline the importance for training procedures of considering which skills are necessary at specific stages in training. That is, it may be crucial for efficient training on this game to concentrate on ship control during the early stages. As training progresses, it may be appropriate to concentrate more on timing of responses.

It is important at this point to consider whether the working memory approach has proved to be any more useful than alternative approaches in the study of the acquisition of complex cognitive skills. A traditional approach to the study of cognitive workload in complex tasks relies on subjective judgements from the trainee of the task or aspects of the task. This can provide useful insights for those involved in training design. However, these judgements should be treated with caution for a number of reasons. For example, the requirement to make subjective judgements about a task may well change the way in which the task is carried out. It is also unclear to what extent individuals have access to the relevant information about their own performance; many aspects of skilled performance may not be available to conscious inspection (e.g. Ericsson and Simon 1980; Nisbett and Wilson 1977). Our own data support this view, since the subjective ratings of difficulty were on the whole poor predictors of performance, particularly with expert players.

A more theoretical approach is to adopt the view that different components of the task compete for general purpose resources. As expertise is gained, some of the processes become relatively automatic and therefore demand less of the available resources (e.g. Shiffrin and Schneider 1977). Our own data suggest that this approach may miss crucial aspects of skilled performance. It is true that general working memory load appeared to have pervasive effects on a number of game components. However, one prediction from the general resources view would suggest that those secondary tasks which were in some sense most difficult should produce the greatest degree of performance decrement when combined with the game. It should also be true that tasks which were by some criteria relatively simple, should produce very little impairment in game performance. In so far as the ratings of difficulty can be used as measures of difficulty, it is clear that these predictions were not supported by our data.

A second difficulty for the general resources view comes from the pattern of disruption shown by different secondary tasks. In experiment 3, our different classes of secondary tasks appeared to affect different clusters of game components. This suggests the involvement of separate specialised resources rather than general resources.

What then do these studies contribute to the development of the concept of working memory? There are strong hints in our data that a working memory in the sense of a set of mechanisms that are involved in monitoring and acting upon several near simultaneous events is indeed useful. This application of the Baddeley and Hitch (1974) working memory model is largely unexplored. There is also evidence that is consistent with the notion of mechanisms that specialise in visuo-spatial processing (Baddeley and Lieberman 1980; Logie 1986; Logie and Baddeley 1989), and in the timing of responses. Further the importance of these mechanisms changes with an increase in expertise on the game. However, the structure and more detailed characteristics of the mechanisms involved remain unclear.

It appears therefore that a number of aspects of these studies are open to investigation. Two sets of studies are immediately apparent. First, these experiments have provided pointers as to the structure of skilled performance, or the acquisition of such skills in terms of subcomponents. It is important to ensure that the results obtained were not simply due to the requirement to perform all of these subcomponents together. One approach to this problem would be to examine whether the patterns of decrement found in the whole game, successfully predict the patterns of disruption on subcomponents of the game when these are performed as part tasks. This would have the important advantage of allowing a more detailed study of the relationship between the various measures of game performance and the underlying cognitive skills that are necessary for particular patterns of performance.

Second, the results of the three experiments reported here have important implications for the framework of working memory in terms of its applicability to rather more complex tasks than hitherto. Thus a second, important area for exploration lies in the generality of this methodology and its applicability to other complex tasks with rather different characteristics, such as greater memory load or more complex and long-term strategic content.

The advantage of these courses would be threefold. They would provide insight into the rather more general question as to the processes which underlie skill acquisition. They would offer a possible new methodology for the measurement of cognitive workload. Finally they would act as a basis for theoretical advances in the development of an increasingly influential and fruitful explanatory framework.

Appendix: Component measures used

In addition to game score, a total of 54 measures were recorded during each game. A complete list of these has been given elsewhere (Mané et al. 1989, this volume). We considered that an analysis based on all of these measures would have been unwieldy and virtually impossible to interpret, as many of the statistical comparisons could have yielded significant differences purely on the basis of chance. Also many of these measures gave unnecessary detail. As such, generally only combinations of two or more measures were included in our analyses. For example, there were separate records of the number of times the ship was damaged by a mine, and the number of times the ship was damaged by the fortress. We considered that it was sufficient to combine these totals as a measure of the subject's ability to avoid damage from either source (NSHDMG). A second example is MOVMNT, which is a combination of separate records for horizontal and vertical movement on the screen. SHTEFF is an example where several measures were combined:

$$\text{SHTEFF} = (\text{NFOHITS} + \text{NMKLFR} = \text{NMKLFO}) \times 100/\text{TOTAL SHOTS}.$$

By this process, we chose seventeen scores that we considered would provide measures of an adequate range of game components. These measures (listed below) formed the basis of the analysis of game microstructure in all three experiments.

MOVMNT: This is a measure of the total amount of movement of the ship around the screen.

NBDBLE: This is a count of the number of times the trigger button on the joystick was pressed quickly, twice in succession, but before ten hits on the fortress.

NBDIFF: This is a count of the number of times the IFF button was pressed when a friendly mine appeared.

NBDINT: This is a count of the number of times the trigger button was pressed twice in succession but where the interpress interval is outside the range of 250–400 msec.

NFODES: This is a count of the number of times that the fortress was destroyed.

NFOHITS:	This counts the number of times the fortress was hit by a missile or a friendly mine
NMKLFR:	This records the number of friend mines energised correctly.
NMKLFO:	This records the number of foe mines successfully destroyed.
NMOVIN:	This counts the number of times that the ship moves into the line of fire of the fortress.
NSHDMG:	This records the number of times that the ship was damaged by the fortress or a mine.
NWRAP:	This is a count of the number of times the ship was moved off the edge of the screen so that it wrapped around to the corresponding position on the opposite side of the screen.
PCNTBON:	This indicates the percentage of occasions when the subject takes advantage of the available bonuses.
RTIFF:	This is the mean reaction time to press the IFF button after a foe mine appears.
SHTEFF:	This is a record of the number of missiles fired by the ship which hit an appropriate target.
TIMKLFO:	This records the mean time to destroy a foe mine from when it appears.
TIMKLFR:	This records the mean time to energise a friendly mine from when it appears.
TOTSHOTS:	This is the total number of missiles fired by the ship.

References

Allport, A., 1980. 'Attention and performance'. In: G. Claxton (ed.), Cognitive psychology: New directions. London: Routledge and Kegan Paul.

Baddeley, A.D., 1979. 'Working memory and reading', in: P.A. Kolers, M.E. Wrolstad and H. Bouma (eds.), Processing of visible language 1. New York: Plenum.

Baddeley, A.D., 1986. Working memory. Oxford, Oxford University Press.

Baddeley, A.D. and G. Hitch, 1974. 'Working memory'. In: G.H. Bower (ed.), The psychology of learning and motivation, Vol. 8. pp. 47–90.

Baddeley, A.D. and K. Lieberman, 1980. 'Spatial working memory'. In: R. Nickerson (ed.). Attention and performance VIII, Hillsdale, NJ: Erlbaum.

Baddeley, A.D., R.H. Logie, S. Bressi, S. Della Sala and H. Spinnler, 1986. Dementia and working memory. Quarterly Journal of Experimental Psychology 38A, 603–618.

Baddeley, A.D., N. Thomson and M. Buchanan, 1975a. Word length and the structure of short-term memory. Journal of Verbal Learning and Verbal Behavior 14, 575–589.

Baddeley, A.D., S. Grant, E. Wight and N. Thomson, 1975b. 'Imagery and visual working memory'. In: P.M.A. Rabbitt and S. Dornic (eds.), Attention and performance V. London: Academic Press. pp. 205–217.

Baddeley, A.D., R.H. Logie, I. Nimmo-Smith and N. Brereton, 1985. Components of fluent reading. Journal of Memory and Language 24, 119–131.

Brooks, L.R., 1967. The suppression of visualization by reading. Quarterly Journal of Experimental Psychology 19, 289–299.

Brooks, L.R., 1968. Spatial and verbal components in the act of recall. Canadian Journal of Psychology 22, 349–368.

Daneman, M. and P.A. Carpenter, 1980. Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior 19, 450–466.
Ericsson, K.A. and H.A. Simon, 1980. Verbal reports as data. Psychological Review 87, 215–251.
Hitch, G.J., 1978. The role of short-term working memory in mental arithmetic. Cognitive Psychology 10, 302–323.
Jonides, J., M. Naveh-Benjamin, and J. Palmer, 1985. Assessing automaticity. Acta Psychologica 60, 157–171.
Keppel, G., 1973. Design and analysis: A researcher's handbook. Engelwood Cliffs, NJ: Prentice-Hall.
Logie, R.H., 1986. Visuo-spatial processing in working memory. Quarterly Journal of Experimental Psychology 38A, 229–247.
Logie, R.H. and A.D. Baddeley, 1987. Cognitive processes in counting. Journal of Experimental Psychology: Learning, Memory and Cognition 13, 310–326.
Logie, R.H. and A.D. Baddeley, 1989. 'Imagery and working memory'. In: P.J. Hampson, D.F. Marks and J.T.E. Richardson (eds.) Imagery: Current developments. London: Routledge and Kegan Paul.
Mané, A.M., J.A. Adams and E. Donchin, 1989. Adaptive and part–whole training in the acquisition of a complex perceptual-motor skill. Acta Psychologica 71, 179–196 (this volume).
Navon, D., 1987. The role of outcome conflict in dual-task interference. Journal of Experimental Psychology: Human Perception and Performance 13, 435–448.
Nisbett, R.E. and T.D. Wilson, 1977. Telling more than you know: Verbal reports on mental processes. Psychological Review 84, 231–259.
Salamé, P. and A.D. Baddeley, 1982. Disruption of short-term memory by unattended speech: Implications for the structure of working memory. Journal of Verbal Learning and Verbal Behavior 21, 150–164.
Shiffrin, R.M. and W. Schneider, 1977. Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. Psychological Review 84, 127–190.
Vallar, G. and A.D. Baddeley, 1982. Short-term forgetting and the articulatory loop. Quarterly Journal of Experimental Psychology 34A, 53–60.
Wickens, C.D., 1980. 'The structure of attentional resources'. In: R.S. Nickerson (ed.), Attention and performance VIII. Hillsdale, NJ: Erlbaum.