Ergonomics
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713701117

Mental performance in extreme environments: results from a performance monitoring study during a 438-day spaceflight

Dietrich Manzey Bernd Lorenz Valeri Poljakov

To cite this Article Poljakov, Dietrich Manzey Bernd Lorenz Valeri(1998) 'Mental performance in extreme environments: results from a performance monitoring study during a 438-day spaceflight', Ergonomics, 41: 4, 537 — 559

To link to this Article: DOI: 10.1080/001401398186991

URL: http://dx.doi.org/10.1080/001401398186991

ERGONOMICS, 1998, VOL. 41, NO. 4, 537–559

Mental performance in extreme environments: results from a performance monitoring study during a 438-day spaceflight

DIETRICH MANZEY†*, BERND LORENZ† and VALERI POLJAKOV‡

†Deutsches Zentrum für Luft- und Raumfahrt (DLR), Department of Aviation and Space Psychology, Institute of Aviation and Space Medicine, Sportallee 54a, D-22335 Hamburg, Germany

‡Institute of Biomedical Problems, Moscow, Russia

Keywords: Spaceflight; Human performance; Cognitive processes; Dual-task; Visuo-motor processes; Subjective workload.

During their stay in a space habitat, astronauts are exposed to many different stressors that may entail detrimental effects on mood and performance. In order to monitor the effects of the space environment on different human information processing functions during an extraordinary long-term space mission, the cognitive, visuo-motor and time-sharing performance of one Russian cosmonaut was repeatedly assessed (29 times) during his 438-day stay in space. The performance tasks used were chosen from the AGARD-STRES battery and included grammatical reasoning, Sternberg memory-search, unstable tracking, and a dual-task consisting of unstable tracking with concurrent memory-search. In addition to performance assessment, several subjective ratings concerning mood and workload were collected. Comparisons of pre-flight, in-flight, post-flight and two follow-up assessments 6 months after the mission revealed, (1) no impairments of basic cognitive functions during the flight, (2) clear impairments of mood, feelings of raised workload, and disturbances of tracking performance and time-sharing during the first 3 weeks in space and the first 2 weeks after return to Earth, (3) an impressive stability of mood and performance during the second to fourteenth month in space, where mood and performance had returned to pre-flight baseline level, and (4) no long-lasting performance deficits at follow-up assessments. From these results it is concluded that the first 3 weeks of long-term spaceflights and the first 2 weeks back on Earth represent critical periods where adverse effects on attentional processes are to be expected, induced by the demands to adjust to the extreme environmental changes. The stability of mood and performance observed after successful adaptation to the space environment indicates that mental efficiency and emotional state can be maintained on a level as high as on Earth even during extraordinary long-term space missions.

1. Introduction

Working under the impact of environmental stressors often compromises human mental efficiency. This has been shown in a large number of laboratory experiments mainly focusing on acute effects of single stressors (e.g. noise, heat, vibration, hypoxia; see for a review Hockey 1986, Smith and Jones 1992). These studies, however, do not allow valid predictions as to how humans can adapt to chronic exposure to conglomerates of superimposed environmental stressors that are usually

*Author for correspondence.

found in what are referred to as extreme environments, i.e. environments to which humans are not naturally suited and which demand complex processes of physiological and psychological adaptation. Outer space represents one example of such an extreme environment. Living and working in a space habitat involves chronic exposure to several environmental stressors that may render it difficult for astronauts to maintain the sufficient level of performance needed for successful operation and crew safety (Kanas 1985, Christensen and Talbot 1986, Luczak 1991, Manzey *et al.* 1995a). Four different kinds of stressors may be differentiated during spaceflights:

(1) Stressors arising from the space environment. The most pronounced example of this kind of stressor is microgravity. Lacking the usual gravitational force induces several physiological changes (e.g. changes of vestibular input; shift of body fluids into the upper parts of the body; changes of bone and mineral metabolism; disturbances of proprioceptive processes) the effects of which have been described in a large number of publications (see for a review Nicogossian *et al.* 1994);
(2) Stressors arising from the space habitat and its life-support system. This class of stressors include factors such as confinement, elevated levels of CO_2 in the ambient air, and elevated noise levels;
(3) Stressors related to the mission-specific physical and mental workload of cosmonauts;
(4) Stressors related to the social situation in the space habitat (e.g. lack of privacy, isolation from family and friends).

Whereas the effects of the latter class of stressors may be expected to emerge only after some weeks in space, all other kinds of stressors may affect the performance and mood of astronauts during both short-term and long-term spaceflights.

Despite its operational relevance, human performance related research during spaceflights, so far, has been limited largely to specific effects of microgravity on visual and perceptual functions, as well as psychomotor processes involved in gross body control (cf. Parker *et al.* 1989). Only in recent years has scientific interest increased to describe more fundamentally the time course of different aspects of mental efficiency throughout space missions. This has led to several performance monitoring studies during short-term spaceflights. Benke *et al.* (1993), for example, assessed the cognitive performance of one cosmonaut during a 6-day space mission to the Russian space station MIR. Using a battery of performance tasks including a simple reaction time task, a choice reaction time task, a Stroop-like interference task, a spatial memory task, and a spatial orientation task they did not find any significant impairments of performance in space. Manzey *et al.* (1993, 1995b) adopted a similar but more extended single-subject approach during an 8-day MIR mission using a subset of the AGARD Standardized Tests for Research with Environmental Stressors (STRES; AGARD 1989). In accordance with the results of Benke *et al.* (1993), cognitive performance, i.e. speed and accuracy of short-term memory retrieval and logical reasoning, was found to be unimpaired during the stay in space. However, performance in an unstable tracking task, demanding fine manual control movements, decreased significantly during the Soyuz approach phase and the last 4 days onboard MIR. In addition, time-sharing efficiency assessed by simultaneous performance of tracking and memory search was found to be impaired throughout the whole stay in space, pointing to possible attentional selectivity effects under the impact of spaceflight-related stressors. Similar results were reported by Schiflett *et al.* (1995) and Schlegel *et al.* (1995), who assessed the performance of three astronauts during a 13-day American shuttle mission. Using very similar tasks to Manzey *et al.* (1993, 1995b) they found clear impairments of tracking and dual-task performance in space. However, deviating from the studies cited above, they also reported some impairments of memory-search speed in two of their astronauts and raised the question of whether their results are microgravity related or present effects of decreased alertness and fatigue.

These results from recent performance monitoring studies suggest that although basic cognitive processes remain more or less stable in space, at least visuo-motor processes and higher attentional functions may be prone to disturbance effects during short-term spaceflights. The present work aims at extending this line of research to long-term spaceflights. Russian psychological support groups distinguish two stages of adaptation to the space environment during prolonged spaceflight (Grigoriev *et al.* 1985). The first stage covers the first 2 to 6 weeks in space, when the cosmonaut has to adapt to the conditions of microgravity and the accompanying physiological changes, as well as to the other environmental conditions in the space habitat. In addition the cosmonaut has to adapt to his new work-rest cycle and the operational demands according to the flight programme. Although empirical data from Russian spaceflights are still lacking in the available literature, it may be assumed that performance decrements due to direct or indirect effects of microgravity or work overload may be observed primarily during this time of the mission. During the second stage of adaptation ('at which the shock of novelty is no longer felt', p.2), emotional effects due to feelings of monotony and the psychosocial conditions during spaceflights may prevail (Grigoriev *et al.* 1985). This matches well with some early results from time and motion studies during long-term Skylab missions lasting 28 to 84 days (Kubis *et al.* 1977). In these studies, clear performance decrements (slowing of motor activity compared to last pre-flight performance) were observed only during the first week in space, when pre-trained tasks had to be executed for the first time under spaceflight conditions.

The recent 438-days space mission of one Russian cosmonaut, which set a new world record for humans in space, provided a unique opportunity to monitor the efficiency of cognitive, visuo-motor and higher attentional processes during an extraordinary long-term space mission. In particular, the following issues were addressed: How long does it take to recover fully from initial performance disturbances during long-term spaceflights? Is it possible to maintain stability of mental efficiency after complete adaptation to space during a long-term spaceflight? Do similar performance decrements as during adaptation to space emerge also during re-adaptation to Earth after a long-term spaceflight? Do prolonged stays in space induce any long-term performance deficits after return to Earth? Performance assessment was done by the same set of laboratory tasks of the AGARD STRES-battery that had been used previously (Manzey *et al.* 1993, 1995b), and included pre-flight, in-flight, and post-flight assessments as well as a follow-up assessment 6 months after the mission. In order to get a comprehensive picture of the mental and emotional state during the stay in space, performance assessment was supplemented by subjective ratings of mood and workload. Concerning the duration of the space mission and the number of performance assessments in space, the results of the present work represent a set of data that for the first time describe in some detail the time-course of performance effects during human adaptation to the space environment, the stability of performance after complete adaptation, and performance-related effects of re-adaptation to Earth conditions after more than 14 months in space.

2. Method

2.1. *Subject*

One Russian cosmonaut (age at beginning of the mission: 51 years, right-handed) participated in the study on his second long-term space mission. His first spaceflight occurred in 1988/1989 and included 241 days onboard MIR.

2.2. *Performance tasks*

Performance tasks were selected from the AGARD battery of Standardized Tests for Research with Environmental Stressors (STRES) (AGARD 1989). Four tasks were used:

(1) *Grammatical Reasoning Task (GRT).* Referring back to a more simple paradigm described by Baddeley (1968), this task required complex logical reasoning operations based on grammatical transformations. Each trial consisted of two statements describing a sequence of three symbols (e.g. & BEFORE *,* AFTER#) which were presented together with a certain set of three symbols (e.g. & * #). The subject had to evaluate whether the two statements were true for the given set of symbols. If the truth values of both statements were the same, the subject had to press a key for 'same'. If the truth values differed (one statement true, the other false) he had to press a key for 'different'. Response times and errors were scored for each single item.

(2) *Memory Search Task (MS2; MS4).* The Sternberg (1966) memory-search task was used to evaluate short-term memory functions. The subject had to memorize a set of letters (the memory-set) and was then presented with a series of single probe letters. By pressing a key for either 'yes' or 'no', he had to indicate whether or not the letter belonged to the memory set. Fixed memory sets of two (MS2) and four (MS4) letters were used in separate 3-min blocks of trials. Response times and errors were scored as the performance measures for each single probe.

(3) *Unstable Tracking Task (UTT).* In this task a horizontally-moving cursor had to be centred by means of a joystick within a marked target located in the middle of the screen. The inherent dynamics of the tracking loop included a positive feedback of the tracking error resulting in system instability which was further increased by a divergent element ($\lambda = 2$) (Jex *et al.* 1966, Allen and Jex 1972). Performance was quantified by calculating the root-mean-square tracking error (RMSE) integrated over blocks of 1 s and averaged across each 3-min run.

(4) *Dual-Task (DT2; DT4).* This task required simultaneous performance of UTT with MS2 (DT2) or MS4 (DT4), respectively, resulting in two versions of dual-task with different memory-load. The subject was instructed to divide his attention equally between both tasks. Performance scores were the same as for the single tasks.

2.3. *Subjective measures*

Subjective measures included ratings of subjective workload experienced during task performance and subjective mood ratings. The NASA Task-Load-Index (TLX, Hart and Staveland 1988) was used for subjective workload assessment. It consists of six different 21-point rating scales that require a subjective evaluation of different workload dimensions (mental demands, physical demands, time pressure, own performance, effort, frustration). For the study, the English scale descriptions were translated into Russian. A thorough instruction of the subject on the use of the rating scales assured that the subject's interpretation of scales did not deviate from the scale meanings in the original NASA-TLX. Mood ratings were recorded by 15, six-point scales (0 – 5) labelled with the following adjectives (in Russian): aggressive, balanced, bored, carefree, concentrated, distracted, fatigued, fresh, happy, interested, lively, nervous, relaxed, sad and strained. A German version of this set of mood scales had already been used before in a 60-day confinement study (Lorenz *et al.* 1996).

2.4. *Apparatus and stimuli*

Presentation of subjective mood and workload scales, performance tasks, as well as response recording and scoring of performance data (response times, tracking error) was controlled by an IBM-compatible laptop (Unisys Powerport 1386 SX, Hamburg, Germany). All performance tasks were generated using a commercially available code generating system (ERTS), Beringer 1993) and were presented on the screen of the laptop, which was positioned approximately 60 cm from the subject. Stimuli of MS2 and MS4 were presented at the screen centre. They were chosen from Windows™ bitmap fonts HELVB with a letter size of 8 mm subtending a visual angle of about 0·7° vertically. The tracking cursor subtended a visual angle of approximately 0·19° horizontally and 1·4° vertically. In the dual-tasks, memory search probes were presented above the target area of the tracking task with a vertical distance of about 2·5° between the target centre and the centre of the probe letter. Responses for MS2/4 and GRT had to be given with the left hand by pressing one of two keys on the keyboard ('D' and 'W'). The tracking task had to be controlled by a joystick, which was located on the right side of the laptop. During the space mission, performance data were stored on floppy discs for off-line data analysis.

2.5. *Procedure*

In order to familiarize the subject with self-administration of the computerized task battery and to reduce the influence of practice on task performance during the experiment, a total of 33 training sessions were performed by the subject over a period of 5 months prior to the space mission. Each training session included one 3-min trial for each single-task and both dual-task versions, and lasted about 30 min. During training, performance feedback (mean of response times for correct responses and number of errors, or mean of RMSE-scores, respectively) was given after each task. The subject was instructed to maximize individual performance by minimizing response times (or tracking error, respectively) while attaining maximum response accuracy without assigning priority to either parameter.

The experiment included 4 pre-flight, 29 in-flight and 8 post-flight assessments. According to the AGARD standard, no performance feedback was given during this phase. Each of the experimental sessions lasted approximately 30 min. At the beginning of each session the cosmonaut completed the mood ratings. After this, the six performance tasks had to be performed for 3 min each in a fixed sequence (GRT, MS2, MS4, UTT, DT2 and DT4) with a subject-controlled intertask interval that differed between 20 s and 1 min. Sequence of tasks was kept fixed across experimental sessions because the inherent limitations of the single-subject repeated-measurement approach in this study did not allow for an experimental control of possible task-order × time-of-measurement interactions. On completion of the performance tasks the cosmonaut received the NASA-TLX items, and was instructed to rate the perceived demands of the task-battery on the different scales. The first two pre-flight baseline data collections (BDC) were performed approximately 3 months and 1 month before launch (launch − 87 and − 34 days, respectively). Two additional pre-flight assessments were conducted at days − 3 and − 2 from launch. Training continued during the period between both BDC series. Inflight data were collected between 8 January 1994 and 22 March 1995 during the entire 438-day stay of the cosmonaut onboard the orbital station MIR. The first eight in-flight assessments were conducted during the first month in space, with the first three assessments at days 4, 5, and 6 in space. Further assessments were performed during the second to fourth month (7 assessments), the seventh to ninth month (7), and the 11th to 14th month (7) in space, respectively. During these assessments the subject worked in a stabilized sitting position with the laptop fixed by a belt-system on his thighs. According to prior experiences (Manzey *et al.* 1993) this procedure guaranteed task performance in a fairly steady and comparable position. Moreover, this procedure made it possible to place the subject in a section of the orbital complex where disturbances by other crew members and/or intercom could be minimized. Post-flight baseline data were collected on day + 4 (2 assessments), + 5, + 6, + 11, and + 12 after landing. Two additional follow-up assessments were conducted approximately 6 months after the mission (day + 168 after landing; time between assessments: 3·5 h).

2.6. *Performance measures and statistical analysis*

Error rates in GRT, MS2 and MS4 were generally low for all sessions and experimental conditions (below 12% in GRT, below 5% in MS2, MS4) and therefore statistical analyses of cognitive task performance were limited to response time data. Prior to analyses each single response time on the cognitive tasks was transformed by a reciprocal transformation into its corresponding response rate. This was done in order to reduce the skewedness usually found in the distribution of response times and to harmonize variances. Analyses of tracking performance were performed on the RMSE-data. Longitudinal changes in the performance data across pre-flight, in-flight, post-flight, and follow-up sessions were evaluated by use of a single-subject design. A strategy was chosen that combined a General Linear Model (GLM) approach to ANOVA with a test on autocorrelated residuals (Hibbs 1974, Revenstorf and Keeser 1989). This approach has already been successfully applied in previous single-subject performance monitoring studies and is described in detail in Lorenz (1994) and Manzey *et al.* (1995b). For each task, the large number of performance scores obtained during the 41 experimental sessions was regarded as an univariate time-series. The basic idea was to estimate the significance of between-session changes in the mean level of the time-series, treating the remaining within-session variance as an estimation of statistical error. Such an approach is justified provided that the time-series shows the entire statistical properties of a white noise process, that is, represents independent sequential observations after removal of all experimental variance. The procedure involves the following steps. First, a set of coding variables specifying the experimental conditions, were generated and – together with the dependent variable – submitted to an ordinary least-squares regression analysis. Second, the autocorrelation function of the regression residuals was computed and judged for the existence of a white noise process (see McCain and McCleary 1979, for a good description of process identification). Particularly, positively autocorrelated residuals cause a substantial inflation of F-scores and thereby lead to an overestimation of the effect under investigation (Hibbs 1974). An appropriate removal of this bias in the significance levels of the regression parameters can be achieved in two ways. The first approach is the computation of a combined regression and a Box-Jenkins time series model. This involves a least-squares fit to the coding variables of the experimental design while simultaneously estimating the Box-Jenkins model parameters by use of non-linear estimation procedures. Residuals of the combined model have again to be examined for the existence of further departures from white noise. A second much more simple way of achieving sequential independency in the data is the establishment of an appropriate time window for data aggregation. According to the authors' experience, in all response rate data that were aggregated over time periods varying between 10 and 20 s, positive autocorrelations found in the ANOVA residuals were caused by non-stationarity within the 3-min trials, which disappeared after incorporating an ordinary least-square linear trend into the experimental design. Therefore, response rate data of MS2/MS4 were averaged across 10-s intervals, GRT data were integrated over 20-s intervals. This procedure reduced the original amount of data of these tasks to 18 (GRT: 9) performance scores for each experimental session. As was confirmed by analyses of autocorrelation functions of residuals, this kind of data aggregation indeed removed any positive autocorrelations from the time-series of residuals for all ANOVA models applied in the analyses of response rates. In the tracking task positive autocorrelation was evident because of the inherent dynamic of the task. Here, data aggregation across 18-s intervals (resulting in 10 performance scores per session) was sufficient to remove this positive autocorrelation. Since an appropriate bias correction was of primary concern the authors preferred the technique of data aggregation as a more pragmatic treatment of the statistical problem of serial-dependent residuals than application of the Box-Jenkins approach. Data aggregation had the further advantage of reducing the vast amount of data.

In the applied GLM ANOVA models, performance data of the 2 pre-flight sessions from days – 87 and – 34 were averaged. This mean was regarded as the baseline measure for evaluation of the experimental session effects. Possible effects of linear trends were controlled by incorporating a covariate in the designs representing the aggregated performance scores for successive time intervals, but statistics of this covariate will not be presented here. The parameter estimates of the regressions for the different levels of the experimental session factor were equivalent to t-tests between mean baseline performances and performance in the respective experimental session and because of the large number of parameters estimated they were treated like two-tailed non-orthogonal *a priori* contrasts with significance levels adjusted according to the Bonferroni procedure ($\alpha = \alpha$/number of comparisons; Kirk 1982: 106ff). Owing to the nature of the data (only one rating per scale and session) and the single-subject approach, no comparable statistical analyses could be performed with the subjective mood and workload ratings.

3. Results

3.1. *Cognitive task performance*

3.1.1. *Grammatical reasoning*: Mean GRT response rates and error rates for pre-flight, in-flight, post-flight sessions, and both follow-up assessments are shown in figure 1. As becomes evident from this figure, grammatical reasoning performance showed a considerable slowing at the two near-launch pre-flight sessions compared to baseline performance 3 months and 4 weeks before launch. During the first 5 days in space, however, a rapid recovery of GRT performance was observed and response rates then remained relatively stable on pre-flight baseline level until post-flight and follow-up sessions. This pattern of effects was confirmed by the ANOVA results which revealed a significant effect of Session ($F(39, 289) = 8{\cdot}61$, $p < 0{\cdot}01$). Pairwise comparisons of baseline performance (mean of performance at days − 87 and − 34) with all other sessions showed significant decrements in response rates only at days − 3, − 2 (both $p < 0{\cdot}01$), and in-flight day 4 ($p < 0{\cdot}03$), and even a significant improvement of performance 6 months after the mission (first assessment at day + 168, $p < 0{\cdot}05$).

3.1.2. *Memory search*: Speed of memory search performance was analysed separately for both levels of memory load by 2-way ANOVAs of response rates with factors defined as Task-Mode (single-task versus dual-task performance) and Session (40 levels). These analyses revealed significant main effects of task-mode for both MS2 ($F(1,1316) = 370{\cdot}40$, $p < 0{\cdot}01$) and MS4 ($F(1,1316) = 159{\cdot}53$, $p < 0{\cdot}01$). For both loading conditions, single-task performance was better than dual-task performance. In addition, the main effects of session became significant (MS2: $F(39,1316) = 70{\cdot}88$, $p < 0{\cdot}01$; MS4: $F(39,1316) = 68{\cdot}44$, $p < 0{\cdot}01$). The corresponding mean response rates for the different sessions (pooled across task-mode factor) and both memory loads are shown in figure 1, together with the mean error rates.

As becomes evident from this figure, the time-courses of memory search performance for pre-flight and first in-flight sessions exhibited entirely the same pattern as the time course of GRT performance. Compared with mean baseline performance at days − 87 and − 34 before launch, speed of memory search declined most clearly at pre-flight days − 3 and − 2, independent of memory-load. A clear recovery of memory search performance was observed already at the first days in space, and memory search response rates even increased above baseline level at several of the following in-flight sessions. The latter effect primarily emerged in MS4 performance. Pairwise comparisons of mean memory search response rates during baseline assessments and the other two pre-flight as well as all in-flight sessions revealed significant performance decrements in MS2 only at days − 3, − 2, 4, 6, 11, and 41 and in MS4 only at days − 3, − 2, 6, and 41 (all $p < 0{\cdot}01$). Significant performance improvements were observed in MS2 at in-flight days 20, 186 (both $p < 0{\cdot}01$), 199 ($p < 0{\cdot}04$), 348, and 420 (both $p < 0{\cdot}01$), and in MS4 at days 27, 34, 48, 68, 186, to 363, 413 and 420 (all $p < 0{\cdot}01$), and 427 ($p < 0{\cdot}04$). In contrast to GRT performance, however, memory search speed showed a clear slowing during re-adaptation after return to earth. Response rates for MS2 and MS4 slowed down during all post-flight assessments (day + 4 to + 12), and — even though a clear recovery from post-flight impairments was observed — remained below pre-flight

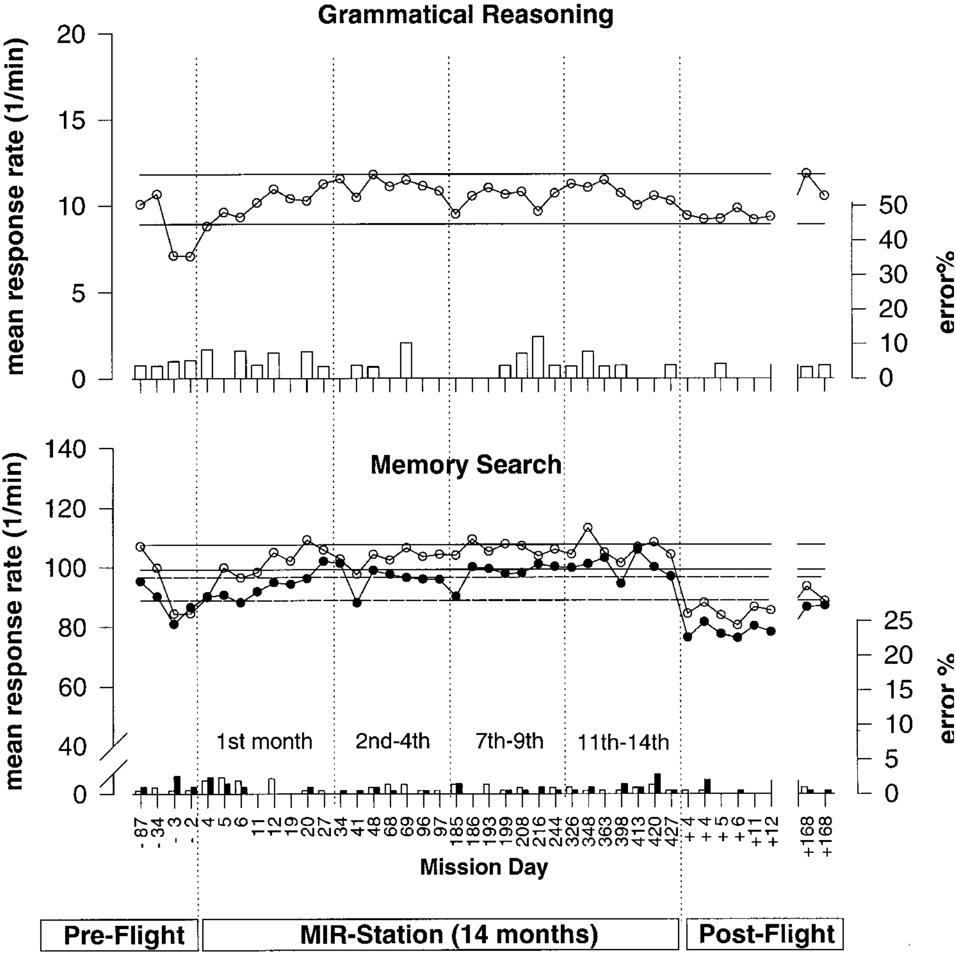

Figure 1. Time-courses of mean response rates (lines) and error rates (bars) as a function of mission-day for the grammatical reasoning task (upper panel) and the memory-search task (pooled across single-task and dual-task performance, lower panel). Memory-search data are presented separately for both levels of memory load: two-letter search (MS2, open circles/bars) and four-letter search (MS4, filled circles/bars). The horizontal lines in the graphs correspond to the upper and lower confidence limits defining mean pre-flight performance at days − 87 and − 34 as the reference for pairwise comparisons (Bonferroni contrasts) with performance at each subsequent session.

baseline level also for the two follow-up assessments (all $p < 0{\cdot}01$). Under both memory loads, the two main effects were qualified by significant Task-Mode × Session interactions (MS2: $F(39,1316) = 9{\cdot}59$, $p < 0{\cdot}01$; MS4: $F(39,1316) = 7{\cdot}20$, $p < 0{\cdot}01$). These interactions indicated that differences between single- and dual-task memory search performance did not remain stable across sessions. The time courses of single- and dual-task response rates for MS2 and MS4 producing these interactions are shown in figure 2, together with corresponding mean error rates. Differences between single- and dual-task performance reflect the costs of dual-task performance, i.e. reflect a slowing in memory-search associated with the demands of concurrent tracking. Bonferroni contrasts revealed that, compared to the mean difference at baseline sessions, these costs increased significantly at in-flight days 5 and 12 for MS2 (both $p < 0.01$) and at days 5, 34 (both $p < 0.01$), and 326 ($p < 0.02$) for MS4. Another increase of the single-dual difference in MS4 performance at day 12 approached significance ($p < 0.10$). These effects were mainly due to a slower recovery of memory search performance under dual- than single-task conditions during the first two weeks in space (figure 2). However, reversed effects, i.e.

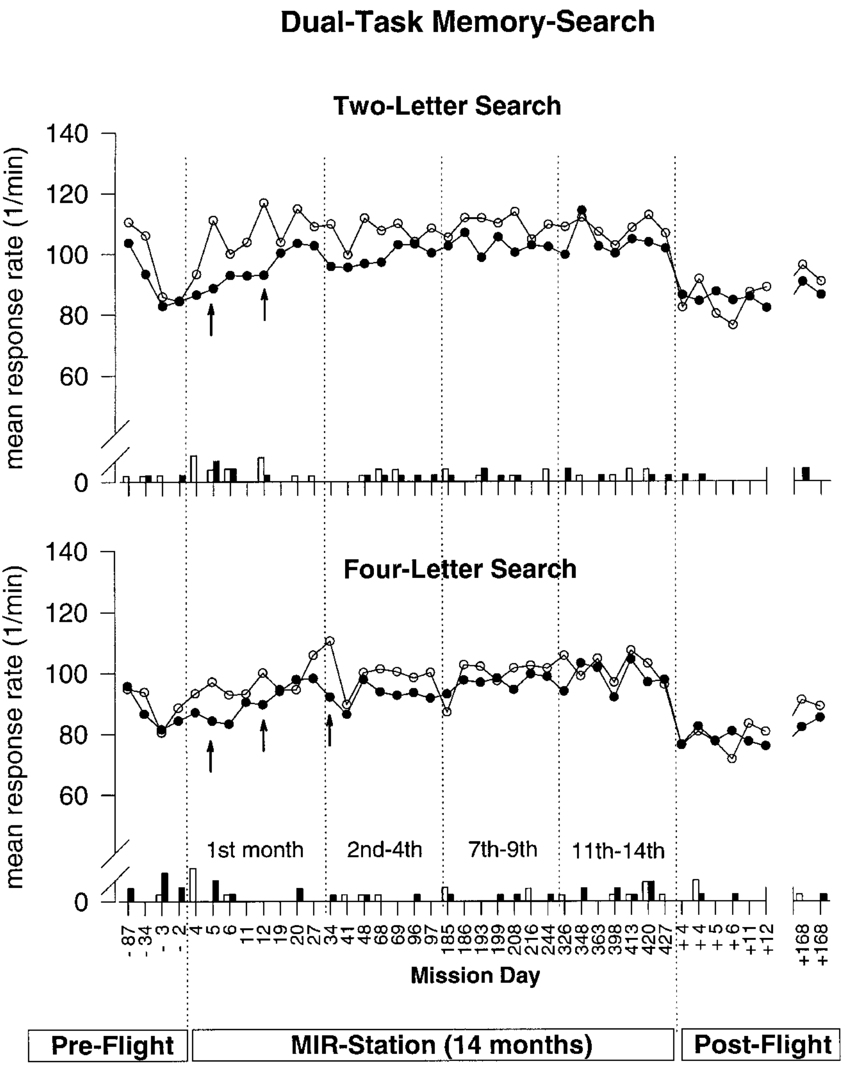

Figure 2. Time-courses of mean response rates (lines) and mean error rates (bars) as a function of mission-day for single-task (open circles/bars) and dual-task memory-search (filled circles/bars). Upper panel: two-letter search. Lower panel: four-letter search. Significant increments of single-dual differences (compared to pre-flight baseline) are marked by arrows.

significantly decreased single-dual differences, were also observed at several experimental sessions. Most striking in this regard were experimental sessions where dual-task performance was better than single-task performance. For MS2 this occurred at pre-flight days − 2, in-flight day 348, and most of the post-flight days (+ 4 [first assessment], + 5, + 6, all $p < 0{\cdot}01$). Another significant decrement of single-dual difference was observed at in-flight day 96 ($p < 0{\cdot}03$). For MS4 dual-task performance was better than single-task performance at in-flight day 185 and post-flight day + 6 (both $p < 0{\cdot}01$). The former effect appeared to be associated with a sharp drop of single-task performance after the long (88 days) break without training on the tasks.

3.2. *Tracking performance*

The tracking data were analysed by a 3 × 40 ANOVA with factors defined as Task-Mode (single-task, dual-task with concurrent two-letter memory search, dual-task with concurrent four-letter search) and Session. This analysis showed that tracking error varied significantly across experimental sessions (main effect of session ($F(39, 1027) = 9{\cdot}05$, $p < 0{\cdot}01$). This main effect is illustrated in figure 3. Most strikingly, tracking error increased during both the first sessions in space as well as the first sessions after return to Earth. Bonferroni contrasts between mean baseline performance and all other sessions revealed significant increases of tracking error at pre-flight day − 3 ($p < 0{\cdot}02$), in-flight days 4, 5, 6, (all $p < 0{\cdot}01$), and 12 ($p < 0{\cdot}02$), and most of the post-flight assessments during the first 2 weeks after landing (days

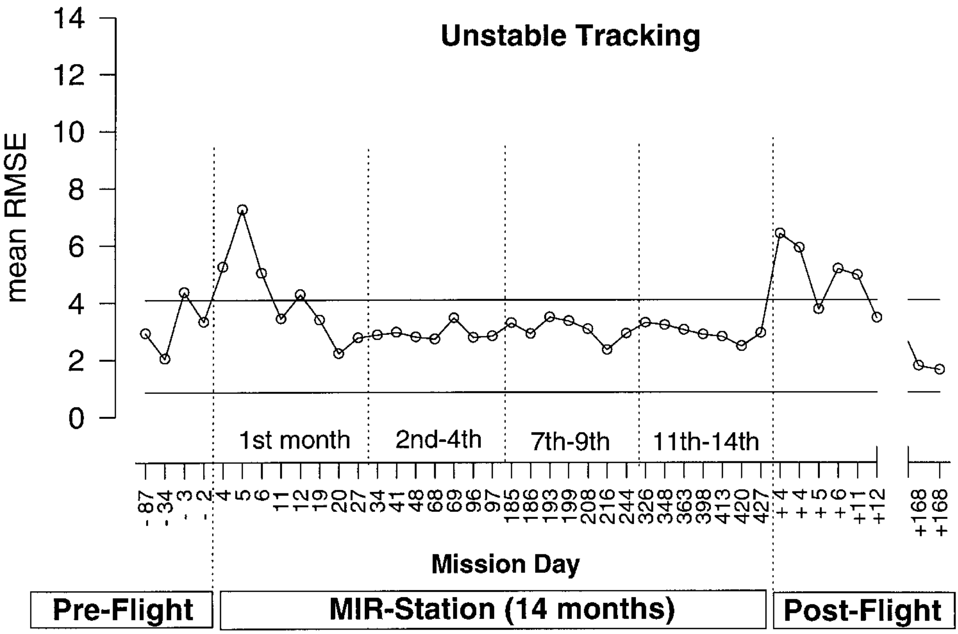

Figure 3. Mean tracking error in the unstable tracking task (pooled across single-task and dual-task conditions) as a function of mission-day. The horizontal lines in the graphs correspond to the upper and lower confidence limits defining mean pre-flight performance at days − 87 and − 34 as the reference for pairwise comparisons (Bonferroni contrasts) with performance at each subsequent session.

+ 4 [both assessments], + 6, + 11, all $p < 0{\cdot}01$). In addition to this effect of session, the main effect of task-mode ($F(2,1027) = 5{\cdot}47$, $p < 0{\cdot}01$) and the Task-Mode × Session interaction ($F(78,1027) = 2{\cdot}09$, $p < 0{\cdot}01$) also became significant. In order to analyse the different effects of task-mode on tracking performance across experimental sessions in more detail, the mean difference in tracking error between dual-task (pooled across memory load conditions) and single-task tracking at baseline days was compared with corresponding differences at all other sessions by Bonferroni *t*-tests (see figure 4 for a contrast of single-task and pooled dual-task tracking performance across sessions). These comparisons revealed that costs of concurrent memory search in tracking increased only at in-flight day 6 ($p < 0{\cdot}07$) and post-flight days + 6 and + 11 ($p < 0{\cdot}06$). A striking reversal of this effect with a clear better dual-task than single-task performance was observed at post-flight day + 4 ($p < 0{\cdot}05$).

3.3. *Time-sharing efficiency*

Although dual-task results have already been described above for memory search and tracking, the separate inspection of results for both tasks does not allow clear conclusions about the time-sharing efficiency of the cosmonaut to be drawn. What is usually observed in dual-tasks is some kind of performance trade-off between both tasks to be combined, i.e. the performance in one task varies dependent on the performance in the other task. Therefore, variations of single-dual differences in one of two combined tasks may not necessarily reflect disturbances of time-sharing efficiency but may only reflect changes in task-emphasis during dual-task

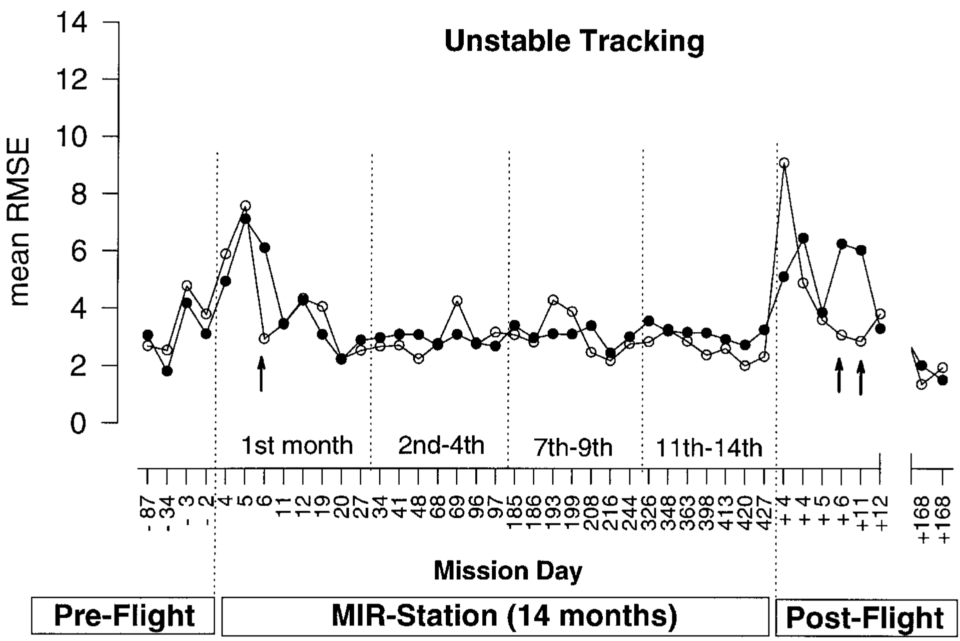

Figure 4. Time courses of mean tracking error for single-task (open circles) and dual-task tracking (filled circles) as a function of mission-day. Dual-task data are pooled across both dual-task conditions. Significant increments of dual-single differences (compared to pre-flight baseline) are marked by arrows.

performance (Navon and Gopher 1979). Consequently, an evaluation of time-sharing efficiency requires that *joint* performance in both combined tasks is considered. Only cases in which increased single-dual differences in one of the tasks are not associated with reversed effects in the other task can be regarded as clear signs of a reduced time-sharing efficiency. In order to analyse this for the dual-task conditions of the present study, the dual-task effects of memory search and tracking are summarized and set against each other in table 1. In this table a plus sign reflects a significant increase of single-dual difference (i.e. increased costs of time-sharing), and a minus sign reflects a significant decrease of single-dual differences (i.e. reduced costs of time-sharing) in the respective tasks. This summary shows that only the dual-task effects observed at post-flight day + 6 appear to reflect a shift of task-emphasis under dual-task conditions (i.e. emphasis shift from tracking to memory search) rather than a real change in time-sharing efficiency. In addition, table 1 shows that increased costs of time-sharing, i.e. drops of time-sharing efficiency, occurred primarily during the first month in space. At later in-flight sessions and most post-flight sessions time-sharing efficiency remained constant at baseline level or showed a significant improvement (most clearly seen at post-flight day + 4 in MS2 and tracking performance where dual-task performance in both tasks was better than single-task performance, cf. figures 2 and 4).

3.4. *Subjective workload*

Subjective ratings of mental, physical, and time-related task demands as well as the invested effort during task-performance are shown for all experimental sessions in figure 5. Ratings of mental demands declined during practice and reached a minimum at day – 34, which might be regarded as baseline rating for this aspect of workload. A similar but weaker trend was also observed for the other ratings. The overall time-course across experimental sessions shows that task demands and the effort needed to perform the different tasks were rated much higher for the first in-flight assessments than for pre-flight assessments. Another sharp increase of demand and effort ratings could be observed during the first 2 weeks after landing. A closer look at the data revealed that the largest increments from pre-flight to first in-flight ratings occurred for ratings of physical demands and effort which both remained clearly above baseline level during the first 3 weeks in space. In contrast, mental task demands were rated higher than baseline already at pre-flight days – 3 and – 2 immediately before launch, but also showed a further increase at in-flight days 4 to 11. Time-pressure ratings raised considerably at days – 3 and – 2 and remained on this elevated level during the first 2 weeks in space. After the first 3 weeks in space all ratings returned to pre-flight baseline level and remained relatively stable on this level throughout the entire stay in space. Only mental demand ratings showed another considerable increase at in-flight day 185 when the tasks had to be performed after a long 88-day break without any training on tasks. After return to Earth, all demand and effort ratings raised again on a level as high as during the first days in space. Mental and physical demand ratings stayed on this elevated level during all post-flight sessions, but were down at baseline values at follow-up assessments. Effort ratings were clearly higher than pre-flight and last in-flight ratings at post-flight days + 4, + 6, + 11, + 12. At follow-up assessments, effort ratings were elevated for the first assessment after almost 6 months without training of the different tasks, but returned immediately to pre-flight baseline level for the second assessment.

Table 1. Summary of dual-task results for unstable tracking and memory search. For each task significant changes of single-dual performance differences (i.e. costs of time-sharing) compared to baseline values are presented (+ : significant increase of single-dual performance difference, – : significant decrease of single-dual performance difference).

	Mission-day												
	Pre-flight	1st month			2nd – 4th months		7th – 9th months	11th – 14th months		Post-flight			
Task	– 2	5	6	12	34	96	185	326	348	+ 4	+ 5	+ 6	+ 11
MS2	–	+		+		–			–	–	–	–	
MS4		+		+	+		–	+				–	
UTT			+							–		+	+

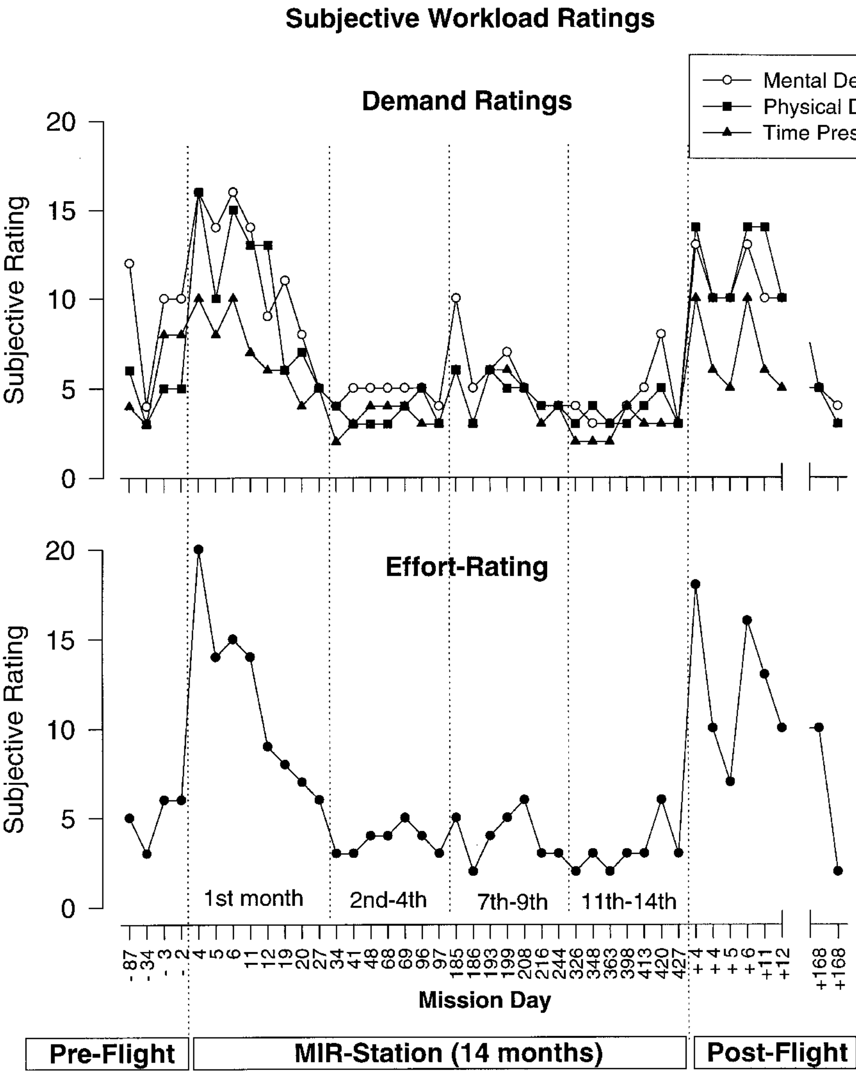

Figure 5. Time-courses of subjective workload ratings as a function of mission-day. Upper panel: Subjective ratings of mental (circles), physical (squares), and time-related (triangles) task-demands. Lower panel: Effort-rating.

3.5. *Subjective mood ratings*

Nine of the 15 mood scales showed extreme low ($< 1{\cdot}0$) or high ($> 4{\cdot}0$) means or only small variations ($s < 1{\cdot}0$) across the 41 experimental sessions: 'aggressive' (mean = 0·05/s = 0·32), 'bored' (0·03/0·16), 'carefree' (0·59/1·02), 'concentrated' (4·51/0·81), 'distracted' (0·78/1·24), 'happy' (0·80/0·75), 'interested' (5·0/0·0), 'nervous' (0·34/0·69) and 'relaxed' (2·24/0·80). Obviously, none of these scales represented aspects of subjective mood that were influenced by the extreme living conditions in space. The other six scales were factor analysed using principal component analysis. This analysis was based on the correlations of the six scales across the 41 experimental sessions. Based on the screen test, two principal components were extracted which represented 59·5% and 16% of the variance in the data, respectively. After Varimax rotation the first (bi-polar) principal component was particularly marked by two scales, 'balanced' (factor loading: 0·88) and 'fatigue' (– 0·77), and therefore represented aspects of subjectively experienced strength (i.e. being balanced and alert). The second principal component almost exclusively represented variations in 'sad' -ratings (factor loading: 0·91). The time-courses of

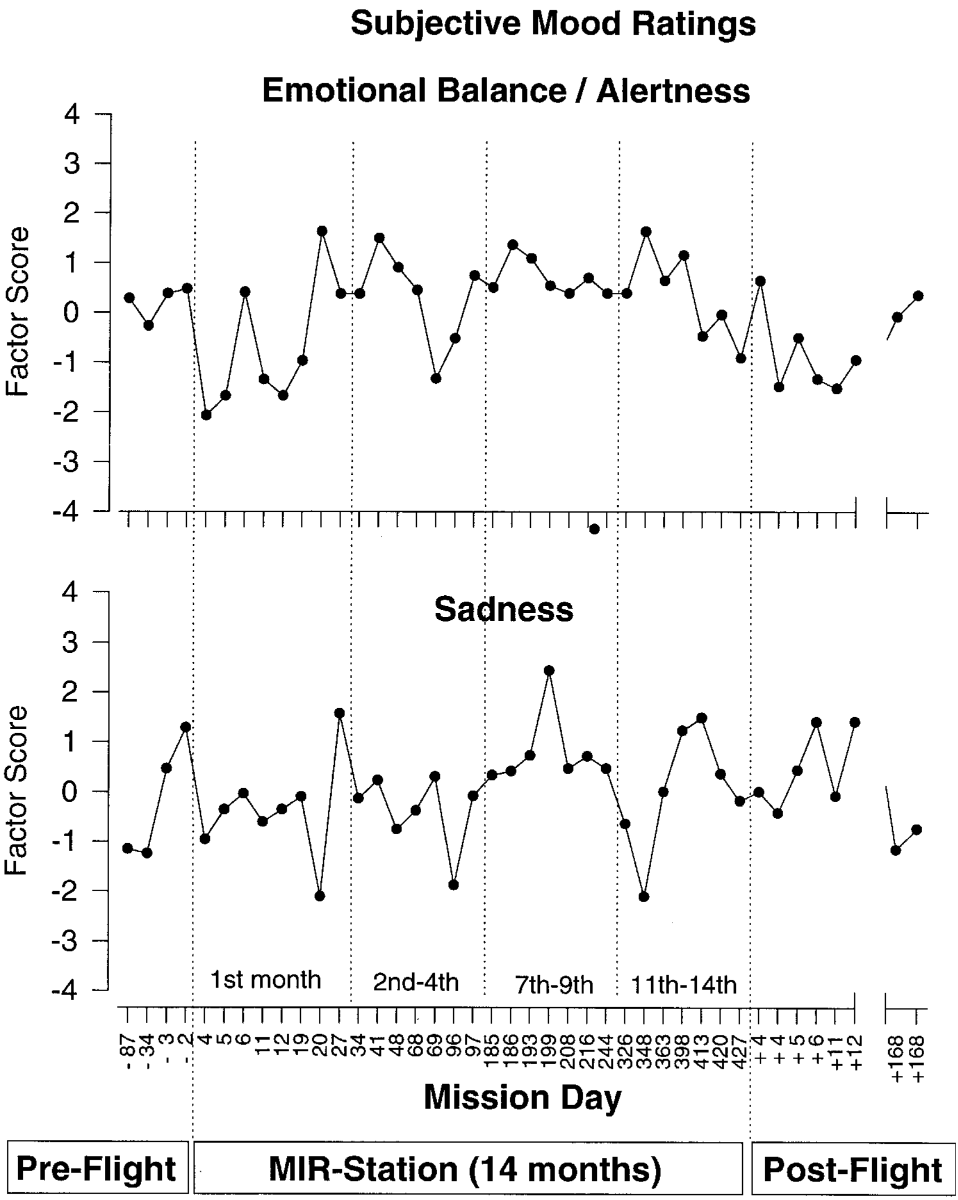

Figure 6. Factor scores for both principal components extracted from subjective mood ratings for pre-flight, in-flight and post-flight assessments. Upper panel: Mood Factor 1 (Emotional Balance/Alertness). Lower panel: Mood Factor 2 ('Sadness').

factor scores of both mood components across experimental sessions are shown in figure 6. As becomes evident from this figure, strength was perceived to be high at all pre-flight sessions, most of the in-flight sessions during the 2nd to 14th month in space, and at follow-up assessments. In contrast to this generally high level of strength, however, considerable changes were observed during the first 3 weeks in space and the first 2 weeks after return to Earth. During these phases perceived strength decreased considerably (most pronounced at in-flight days 4, 5, 11, 12, 19 and post-flight days + 4 [second assessment], + 6, + 11, + 12).

'Sadness' showed a striking increase immediately before launch, compared to pre-flight baseline values at days − 87, − 34 and both follow-up assessments, but was back to baseline level already at the first assessment in space. During the following months in space, rated sadness increased slightly above baseline values at most of the in-flight days, with obvious deviations from this general level at days 20, 96, 348 (lowest scores) and days 27, 199, 398, 413 (highest scores). After return to Earth sadness ratings remained on the general in-flight level at the first three assessments, and were clearly elevated at post-flight days + 6 and + 12.

In order to identify possible covariations between mood changes and performance effects, bivariate product-moment-correlations were calculated between the factor scores for both mood components and the different performance scores across all 41 experimental sessions. These analyses revealed significant correlations only between the first mood factor (strength) and both single-task tracking performance as well as all dual-task performance scores. Correlations between the 'strength'-component and tracking error were $r = -0{\cdot}34$ (UTT, $p < 0{\cdot}04$), $r = -0{\cdot}54$ (DT2, $p < 0{\cdot}01$), and $r = -0{\cdot}40$ (DT4, $p < 0{\cdot}01$). The corresponding correlations for dual-task memory-search response rates were $r = 0{\cdot}49$ (DT2, $p < 0{\cdot}01$) and $r = 0{\cdot}42$ (DT4, $p < 0{\cdot}01$). No correlations were found between the mood factors and single-task cognitive performance (GRT, MS2, MS4).

4. Discussion

The results of the present study provide the first insights into human performance efficiency during extraordinarily prolonged spaceflights. First, as compared to baseline performance 4 weeks and 3 months before the mission, performance decrements were observed at the last 3 days before launch. This effect most clearly was reflected in a significant slowing of response rates for both cognitive tasks, but was also observable in tracking performance. Second, two phases of de-stabilization could be identified that coincided with adaptation to the space environment (first 3 weeks in space) and re-adaptation to Earth conditions after the flight. These phases were associated with both considerable decrements of tracking performance, as well as elevated workload ratings and clear drops in subjective mood. In addition, occasional impairments of time-sharing efficiency were found. Third, an impressive stability of mood and performance was observed during the second to fourteenth month in space. Fourth, no obvious long-lasting performance deficits were observed after the mission.

The clear decline of response rates for both cognitive tasks, GRT and MST, at days − 3 and − 2 from launch, and the slight increase of tracking error at day − 3 point to a general slowing of cognitive processing speed at this time of the mission. This effect could only be detected as a result of a first baseline data collection well in advance of the launch and has not been reported in previous studies (Benke *et al.* 1993, Manzey *et al.* 1993, 1995b). It might be related to effects of raised mental workload, reflected in the increased subjective ratings of mental demands, time-pressure and effort at these pre-launch days. However, such an hypothesis is contradicted by the rapid recovery of cognitive performance during the first week in space, where demand and effort ratings still increased and remained on a comparatively high level. More likely, therefore, the pre-launch slowing of processing speed might be attributed to an increased emotional load of the cosmonaut facing the launch of this extraordinarily long space mission, which has not been attempted by a human being before. This interpretation gets support from the increase of 'sadness'-ratings observed at pre-flight days − 3 and − 2.

Most important for our understanding of human performance efficiency during adaptation to extreme environmental changes are the mood and performance results obtained for the first 3 weeks in space and the first 2 weeks after return to Earth. Both of these phases represent critical adaptational phases that are associated with drastic physiological and psychological changes in the cosmonaut induced by changes of gravity and general living conditions. In the present study these phases were marked by subjective feelings of reduced strength (i.e. reduced alertness and balance), raised workload, and clear disturbances of tracking performance. In addition, also occasional drops of time-sharing efficiency were observed.

The impairments of subjective mood and elevated workload ratings may be attributed to both the impact of adaptational demands on subjective well-being, as well as some additional effort required by the execution of pre-trained tasks under the new environmental conditions. Although no statistical analyses were performed on the subjective data, the effects are considered to be reliable. This is concluded from the strength and striking time course of effects, i.e. the considerable changes of mood and workload ratings in both adaptational phases compared to the stability of ratings across pre-flight and later in-flight sessions.

With regard to performance effects, particularly the results for the tracking task merit discussion. Whereas both cognitive tasks showed a rapid recovery from pre-launch disturbances during the first days in space, tracking performance declined further, remained worse than pre-flight baseline level for the first three assessments in space, and did not show a complete re-stabilization before in-flight day 20. After return to Earth, i.e. after 14 months in space and complete adaptation to the microgravity environment, tracking performance declined again, most obviously seen at the first two assessments on Earth (post-flight day + 4).

The striking dissociation of tracking performance and performance of both cognitive tasks during the first 20 days of the spaceflight is in accordance with previous results from short-term space missions (Schiflett *et al.* 1995, Manzey *et al.* 1993, 1995b). Using entirely the same performance task battery as in the present study, Manzey *et al.* (1993) found performance in the unstable tracking task impaired during the entire period of an 8-day space mission, whereas grammatical reasoning and memory search performance remained unaffected in space. Similar to the results of the present study, most severe impairments of tracking performance emerged during the first four assessments in space. Shiflett *et al.* (1995) reported preliminary results from a 13-day American shuttle-flight, which also showed worse performance decrements in a (critical) tracking task and only minor variations in several cognitive tasks. However, due to the limited duration of the spaceflights the results of these previous studies were ambiguous with regard to the time-course of the phenomenon, i.e. whether the disturbances of visuo-motor performance observed in tracking does reflect a temporary or chronic performance disturbance during spaceflights. The results of the present study suggest that disturbances of visuomotor performance only emerge during the first 3 weeks in space. After this period, performance returns to pre-flight baseline level and remains on this level without any larger variations, even if the exposure to the space conditions is as long as in the present study and even if there are long intervals during this period without any training on task (compare in-flight tracking performances at days 97 and 185). Transient performance decrements during long-term spaceflights have already been reported from Skylab missions (Kubis *et al.* 1977). These experiments included time and motion studies of activities associated with preparation and execution of medical and scientific experiments, and revealed a clear slowing of performance during the first execution of these activities in space which usually took place within the first week of the spaceflight. Similar to the present study, these effects were most pronounced for fine motor activity. A full recovery from impairments was observed during the second week in space, when the tasks had to be executed for the second time.

Comparable post-flight effects of (tracking) performance, mood and subjective workload have not been reported so far (Benke *et al.* 1993, Manzey *et al.* 1993, 1995b, Schiflett *et al.* 1995). Thus, these effects appear to be specifically related to long-term space missions that allow for a complete adaptation of cosmonauts to space conditions and, therefore, require a more demanding re-adaptation after return to Earth than short-term space missions.

Given the myriad of stressors affecting a cosmonaut during adaptation to space and re-adaptation to Earth after a long-term spaceflight, and given the complexity of the unstable tracking task, which demands both psychomotor and attentional processes, a clear-cut interpretation of the observed impairments of tracking performance is difficult. In a previous paper (Manzey *et al.* 1993) it was hypothesized that tracking impairments in space are due to combined effects of decreased alertness, which might adversely affect attentional processes, and microgravity-related disturbances of sensorimotor processes, with the latter affecting tracking only at the first assessments under microgravity. Similarly, Schiflett *et al.* (1995) supposed that raised fatigue resulting from adaptational demands and workload would be one of the main reasons for (tracking) performance decrements during short-term spaceflights. The results of the present study provide more evidence for this assumption. During the first 3 weeks in space, decrements of tracking performance were clearly correlated with subjective feelings of reduced emotional balance and raised fatigue reflected in low 'strength'-component scores, and a corresponding relation was also found during the post-flight period. The fact that a similar relationship did not emerge for the cognitive tasks is in accord with other studies that have repeatedly shown that tracking performance — due to its attentional demands — is particularly sensitive to effects of fatigue (Batejat and Lagarde 1992, Lorenz *et al.* 1996). Thus, the tracking effects observed in the present study appear to reflect primarily disturbances of attentional processes, induced by adverse effects on subjective well-being and alertness during the demanding adaptational phases.

Even though this might explain most of the tracking decrements, the striking correspondence of largest tracking impairments at the very first in-flight and post-flight assessments to changes in gravity conditions at least suggest the possibility that also disturbances of sensorimotor processes during adaptation to microgravity and re-adaptation to 1 *g* conditions might have contributed to these effects. This assumption gets some support from previous experiments during spaceflights and ground-based studies, which demonstrate that changes in gravity conditions may require adjustments of existing central motor programmes and may cause substantial changes of proprioceptive processes associated with execution of precise voluntary movements (Bock *et al.* 1992, Kozlovskaya *et al.* 1993, Bock 1994). It might be supposed that such effects also have affected the motor-related processes of tracking in the present study, and have required (at least to some extent) a re-learning of tracking at the initial in-flight and post-flight phase. However, the data of the present study do not warrant any clear conclusions in this regard. In addition, research about adjustments of central motor programmes and distortions of proprioception under changed gravity, so far, has largely been limited to aimed arm movements and gross body control, and it remains to be shown whether similar effects indeed may affect fine manual control movements in the horizontal plane like the ones involved in tracking tasks.

The occasional impairments of time-sharing efficiency, which were observed primarily during the first month in space, were reflected in raised single- versus dual-task differences of either memory-search speed or tracking error. The nature of the dual-task used and the independence of these effects from memory load caused them to appear to reflect effects of raised peripheral interference (Wickens 1984) between both visual tasks due to attentional selectivity effects (Easterbrook 1959, Hockey and Hamilton 1983). This has been discussed at length elsewhere (Manzey *et al.* 1995b), and is not repeated here. Easterbrook (1959) has defined 'attentional selectivity' as a reduction in the range of cues that can be attended to simultaneously and supposes that it accompanies states of high emotional arousal. According to Hockey (1986, Hockey and Hamilton 1983) raised attentional selectivity represents one of the most pronounced indicators of stress states induced by several environmental stressors, and has also been found to accompany states of decreased alertness. Thus, impairments of time-sharing efficiency may be regarded as another transient behavioural effect associated with the overall demands put on the cosmonaut during adaptation to the extreme living conditions in space. This conclusion is in accord with previous results (Manzey *et al.* 1995b) and is also supported by high correlations between all dual-task performance scores and subjective ratings of emotional balance and fatigue. A comparable pattern was not found with any of the single-tasks.

One last result of the present study that should be discussed is the slowing of memory-search response rates after the spaceflight. Whereas tracking decrements and dual-task interference effects emerged during both critical adaptational phases, a clear decline of memory-search speed was found only at post-flight sessions. This effect is difficult to explain and its origin remains unclear. Given the rapid recovery from pre-launch impairments during the first days of the spaceflight and the comparatively high memory-search performance level during the following months in space, which prove that memory-search performance neither was affected by the adaptational processes at the beginning of the mission nor by any other factors during the long-term stay onboard MIR, it can be excluded that the post-flight decrement of memory-search response rates represents an adverse after-effect of the extraordinarily long spaceflight. Similarly, it is unlikely that this effect reflects a general slowing of memory-processes caused by the demands of re-entry and re-adaptation after the flight. Such an interpretation not only would be in conflict with the comparatively good memory-search performance during the similar demanding adaptational processes during the first days in space, but would also contradict the unimpaired GRT performance at post-flight sessions. In any case, the clear recovery of memory-search response rates 6 months after the mission and the return of UTT performance and time-sharing efficiency to pre-flight baseline level at follow-up assessments, provide strong evidence that even extraordinarily long space missions do not lead to long-lasting impairments of higher mental functions.

In summary, the results of the present study show that long-term space missions may be associated with disturbances of attentional processes during adaptation to living conditions in space and re-adaptation to Earth conditions after the flight. These effects appear to be primarily related to general effects of adaptation and re-adaptation induced by the conglomerate of stressors that affect the cosmonaut during these critical phases and which also are reflected in subjective feelings of reduced personal strength and elevated workload. To what extent also specific effects of the environmental changes (i.e. changes of gravity conditions), which may be expected to affect sensorimotor processes, have contributed to these effects (in particular the decrements of tracking performance) remains an open question for further research. The time course of performance decrements, feelings of raised workload and impairments of subjective mood observed in the present study suggests that the critical adaptational phase in space includes the first 3 weeks of a long-term space mission. This agrees very well with experiences from Russian psychological support groups who differentiate two stages of adaptation to spaceflights with the first stage of primary adaptation lasting 2 to 6 weeks (Grigoriev *et al.* 1985). The complete recovery from performance and mood disturbances after the first 3 weeks in space and the stability of mood and performance during the following months show that — after successful adaptation to the space environment — it is possible to maintain mental efficiency on a comparatively high level even during extraordinarily long-term space missions. Finally, the results of follow-up assessments 6 months after the mission reveal that long-lasting performance disturbances are not expected after long-term spaceflights.

In evaluating these conclusions it must be taken into account that they are based on single-case data. Since no other astronaut/cosmonaut, so far, has spent a comparatively long period in space, comparable data from other studies are not available yet. However, the convergence of results from the first weeks in space with previous results from similar performance monitoring studies during short-term spaceflights is striking and suggests some generalizability of results. This may not necessarily hold as well for the conclusions concerning the stability of performance efficiency during long-term space missions and the lack of long-lasting performance disturbances after such flights. These conclusions should be regarded as preliminary until they have been cross-validated by further performance monitoring studies during long-term space missions.

Acknowledgements

Many thanks are due to Jürgen Drescher and Alexander Gundel (both DLR) as well as Vladimir Nalishiti and Yuri Shpatenko from Yuri Gagarin's Cosmonauts Training Centre in Starcity, Russia. Without their invaluable help and support it would not have been possible to conduct the present study.

References

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT (AGARD). 1989, *Human Performance Assessment Methods (AGARDograph No. 308)* (Neully Sur Seine: AGARD).

ALLEN, R. W. and JEX, H. R. 1972, Visual-motor response of crewman during a simulated 90-day space mission as measured by the critical task battery, *Proceedings of the 7th Annual Conference on Manual Control, NASA SP-281* (Washington: NASA), 239 – 246.

BADDELEY, A. D. 1968, A 3-minute reasoning test based on grammatical transformations, *Psychonomic Science*, **10,** 341 – 342.

BATEJAT, D. and LAGARDE, D. 1992, Circadian rhythm and sleep deprivation: effects on psychomotor performance, *Medical Science Research*, **20,** 167 – 168.

BENKE, T., KOSERENKO, O., WATSON, N. V. and GERSTENBRAND, F. 1993, Space and cognition: the measurement of behavioral functions during a six-day space mission, *Aviation, Space, and Environmental Medicine*, **64,** 376 – 379.

BERINGER, J. 1993, *Development of a Code Generating Software for Conducting Psychological Reaction Time Experiments* (in German) (Frankfurt/Main: Peter Lang).

BOCK, O. 1994, Joint position sense in simulated changed-gravity environments, *Aviation, Space, and Environmental Medicine*, **65,** 621 – 626.

BOCK, O., HOWARD, I. P., MONEY, K. E. and ARNOLD, K. E. 1992, Accuracy of aimed arm movements in changed gravity, *Aviation, Space, and Environmental Medicine*, **63,** 994 – 998.

CHRISTENSEN, J. M. and TALBOT, J. M. 1986, A review of the psychological aspects of space flight, *Aviation, Space, and Environmental Medicine*, **57,** 203 – 212.

EASTERBROOK, J. A. 1959, The effect of emotion on cue utilization and the organization of behavior, *Psychological Review*, **66,** 183 – 201.

GRIGORIEV, A. I., KOZERENKO, O. P. and MYASNIKOV, V. I. 1985, Selected problems of psychological support of prolonged space flights, Paper presented to the 36th IAF Congress, 7 – 11 October, Stockholm, Sweden.

HART, F. G. and STAVELAND, L. E. 1988, Development of NASA-TLX (Task Load Index): results of empirical and theoretical research, in P.A. Hancock and N. Meshkati (eds), *Human Mental Workload* (Amsterdam: North-Holland), 139 – 183.

HIBBS, D. A. 1974, Problems of statistical estimation and causal interference in time series regression models, in H.A. Costner (ed.), *Sociological Methodology 1973 – 1974* (San Francisco: Jossey-Bass), 252 – 308.

HOCKEY, G. R. J. 1986, Changes in operator efficiency as a function of environmental stress, fatigue, and circadian rhythms, in K.R. Boff, L. Kaufman and J.P. Thomas (eds), *Handbook of Perception and Human Performance*, vol. II. (New York: Wiley), 44.1 – 44.49.

HOCKEY, G. R. J. and HAMILTON, P. 1983, The cognitive patterning of stress states, in G. R. J. Hockey (ed.), *Stress and Fatigue in Human Performance* (New York: Wiley), 331 – 362.

JEX, H. R., MCDONNELL, J. D. and PHATAK, A. V. 1966, A 'critical' tracking task for manual control research, *IEEE Transactions on Human Factors in Electronics*, **HFE-7** 138 – 144.

KANAS, N. 1985, Psychosocial factors affecting simulated and actual space missions, *Aviation, Space, and Environmental Medicine*, **56,** 806 – 811.

KIRK, R. E. 1982, *Experimental Design*, 2nd ed. (Belmont: Brooks/Cole Publishing Company).

KOZLOVSKAYA, I. B., BURLACHKOVA, N. I., GANCHEV, G., GATEV, P., GERSTENBRAND, F. and BERGER, M. 1993, Mechanisms of sensory-motor adaptation to weightlessness. Paper presented to the 10th IAA Man in Space Symposium, Tokyo, 19 – 23 April. (Abstract).

KUBIS, J. F., MCLAUGHLIN, E. J., JACKSON, J. M., RUSNAK, R., MCBRIDE, G. H. and SAXON, S. V. 1977, Task and work performance on Skylab missions 2, 3, and 4: time and motion study—experiment M151, *Biomedical Results from Skylab, NASA SP-377* (Washington: NASA), 136 – 154.

LORENZ, B. 1994, Cognitive and psychomotor performance of divers under simulated deep-sea conditions: methodology and results of human performance monitoring based on single-case analyses (DLR-FB 94-07, in German), Technical report, DLR: Cologne.

LORENZ, B., LORENZ, J. and MANZEY, D. 1996, Performance and brain electrical activity during prolonged confinement, in S.L. Bonting (ed.), *Advances in Space Biology and Medicine*, Vol. 5 (Greenwich: JAI), 157 – 183.

Luczak, H. 1991, Work under extreme conditions, *Ergonomics*, **34,** 687 – 720.

Manzey, D., Lorenz, B., Schiewe, A., Finell, G. and Thiele, G. 1993, Behavioral aspects of human adaptation to space: analyses of cognitive and psychomotor performance in space during an 8-day space mission, *Clinical Investigator*, **71,** 725 – 731.

Manzey, D., Lorenz, B., Schiewe, A., Finell, G. and Thiele, G. 1995b, Dual-task performance in space: results from a single case study during a short-term space mission, *Human Factors*, **37,** 667 – 681.

Manzey, D., Schiewe, A. and Fassbender, C. 1995a, Psychological countermeasures for extended manned space flight, *Acta Astronautica*, **35,** 339 – 361.

McCain, L. J. and McCleary, R. 1979, The statistical analysis of the simple interrupted time-series quasi-experiment, In T.D. Cook and D.T. Campbell (eds), *Quasi-Experimentation* (Chicago: Rand-McNally).

Navon, D. and Gopher, D. 1979, On the economy of the human processing system, *Psychological Review*, **86,** 214 – 255.

Nicogossian, A. E., Huntoon, C. L. and Pool, S. L. (eds) 1989, *Space Physiology and Medicine*, 2nd edn. (Philadelphia & London: Lea and Febiger).

Parker, P. E., Reschke, M. F. and Aldrich, N. G. 1989, Performance, in A.E. Nicogossian, C.L. Huntoon and S.L. Pool (ed.), *Space Physiology and Medicine*, 2nd edn. (Philadelphia & London: Lea and Febiger), 167 – 178.

Revenstorf, D. and Keeser, W. 1989, Zeitreihenanalyse von Therapieverläufen—ein Überblick, in F. Petermann (ed.), *Einzelfallanalyse* (München: Oldenbourg), 167 – 212.

Schiflett, S., Eddy, D., Schlegel, R., French, J. and Shehab, R. 1995, Performance Assessment Workstation (PAWS). Unpublished Final Science Report to NASA (Marshall Space Flight Center: NASA).

Schlegel, R., Shehab, R., Schiflett, S., French, J. and Eddy, D. 1995, The NASA performance assessment workstation: Astronauts vs. a ground-based reference group. Paper presented to the 11th Man in Space Symposium, Toulouse, 27 – 31 March (Abstract).

Smith, A. P. and Jones, D. M. 1992, *Handbook of Human Performance*, Vol. 1: *The Physical Environment* (London: Academic Press).

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

Wickens, C. D. 1984, Processing resources in attention, in R. Parasuraman and D.R. Davies (eds), *Varieties of attention* (Orlando: Academic Press), 63 – 102.