Work content influences on cognitive task load, emotional state and performance during a simulated 520-days’ Mars mission

Iris Cohen\textsuperscript{a,b,*}, Nadia den Braber\textsuperscript{b}, Nanja J.J.M. Smets\textsuperscript{a,b}, Jurriaan van Diggelen\textsuperscript{b}, Willem-Paul Brinkman\textsuperscript{a,**}, Mark A. Neerincx\textsuperscript{a,b}

\textsuperscript{a}TNO, P.O. Box 23, 3769 ZG, Soesterberg, The Netherlands
\textsuperscript{b}Delft University of Technology, Mekelweg 4, 2628 CD, Delft, The Netherlands

A B S T R A C T

In high-risk domains such as human space flight, cognitive performances can be negatively affected by emotional responses to events and conditions in their working environment (e.g., isolation and health incidents). The Cognitive Performance and Error (COPE) model distinguishes effects of work content on cognitive task load and emotional state, and their effect on the professional’s performance. This paper examines the relationships between these variables for a simulated Mars-mission. Six volunteers (well-educated and -motivated men) were isolated for 520 days in a simulated spacecraft in which they had to execute a (virtual) mission to Mars. As part of this mission, every other week, several computer tasks were performed. These tasks consisted of a negotiation game, a chat-based learning activity and an entertainment game. Before and after these tasks, and after post-task questionnaires, the participants rated their emotional state consisting of arousal, valence and dominance, and their cognitive task load consisting of level of information processing, time occupied and task-set switches. Results revealed significant differences between cognitive task load and emotional state levels when work content varied. Significant regression models were also found that could explain variation in task performance. These findings contribute to the validation of the COPE model and suggest that differences in appraisals for tasks may bring about different emotional states and task performances.

1. Introduction

Different professionals, such as police officers, military personnel, pilots and astronauts, occasionally enter high-risk situations, in which the risk for harm, information uncertainty and time pressure evoke stress in the professionals involved (Driskell & Johnston, 1998). Their job is to remain focused and perform well in these situations. Extreme levels of stress, however, can affect cognitive performances in negative ways and consequently deteriorate performances (Keinan, Friedland, & Ben-Porath, 1987; Ozel, 2001; Starcke & Brand, 2012).

Insight into human and work content factors that determine cognitive task performance in these situations are useful for finding ways to counteract the performance decline. When the influences of these factors are known, the focus of support can be placed where the help is needed. It might also allow for better anticipation for such situations (e.g., an improved human resource deployment). By monitoring the human and content variables that affect task performance, content-sensitive and personalized task support can be provided.

Based on a literature study and domain analyses, Cohen, Brinkman, and Neerincx (2012) proposed the Cognitive Performance and Error (COPE) model as a general foundation for task support in high-risk domains. In several empirical studies, this model was refined, “parameterized” and evaluated for different application domains. This paper studies the influences from different work contents on core variables of the COPE model (i.e. cognitive task load and emotional state) and the prediction of task performance based on these variables. The analysis centres around a unique experiment on human space flight: the Mars500 program\footnote{www.esa.int/Our_Activities/Human_Spaceflight/Mars500.}.
COPE-model provide such a basis. The COPE-model of the next section might continue discussing factors that affect performances during long-term isolation missions. The Mission Execution Crew Assistant (MECA) is developing personal ePartners that regularly monitor crew-members cognitive task load and emotional states during individual and joint task performances overall mission phases (M.A. Neerincx et al., 2008). This monitoring is a joint crew-ePartner activity and the basis of envisioned ePartner support functions that should help to better cope with the social, cognitive and affective burdens mentioned above (Diggelen & Neerincx, 2010; Gorbunov, Barakova, Ahn, & Rauterberg, 2011; Hennes, Tuyls, Neerincx, & Rauterberg, 2009). The COPE-model of the next section might provide such a basis.

1.1. COPE-model

Fig. 1 shows a graphical representation of the COPE model which represents the influence of acute stress on working performances (Cohen et al., 2012). It consists of three components: work content, cognitive and affective factors, and the actions. Models have been proposed showing similar relationships between environment, appraisal and performance (Hart & Staveland, 1988; Salas, Driskell, & Hughes, 1996), the COPE model, however, includes measures of objective stress, such as physical measures, to assess and predict performance, instead of focussing merely on subjective levels and measures (Robert & Hockey, 1997; Sanders, 1983).

The COPE model distinguishes work content aspects that influence the performance under stress. The specific task goals and task demands of the work will characterize the involved cognitive and affective processes (H. J. Veltman & Jansen, 2004; J. A. Veltman & Jansen, 2003). When an individual perceives a task (i.e., the demands and goal), an assessment is made that leads to the appraisal of the task as either a threat or a challenge (Lazarus, 1999), and a level of perceived task demand which can deviate from the ‘regular’ task demand level. Goals, often structured in a hierarchical way, drive the performance, but may be appraised differently (e.g., due to its relevance for a higher order goal). More challenging goals improve performance compared to easy goals (Locke, Shaw, Saari, & Latham, 1981). Tasks with different structures or characteristics might also provoke different goals and thereby show differences in task performance.

The task demands need to be met to complete the task successfully. Simple tasks have low task demands; more complex tasks have higher task demands. In the COPE model, (perceived) task demand corresponds to the Cognitive Task Load (CTL) measures (M. A. Neerincx, 2003). The CTL model distinguishes three load dimensions: time occupied (TOC), level of information processing (LIP), and task set switches (TSS). TOC is the fraction of the time that is actually needed to complete the task and the time that is available to complete the task. TSS is a measure for switching between tasks; in complex situations, multiple tasks need to be performed at the same time. It takes attention and effort to complete one task, and activate (i.e., start or continue with) the next. LIP is based on the levels of cognitive processes by Rasmussen (1982) and dual process theories (e.g. (Evans, 2003)). Cognitive processes can be distinguished on a continuum from analytical to intuitively. Whether someone's cognitive processing leans more towards one or the other depends, according to Hammond (1988), on the failure or success of previous judgments, and on the task characteristics. As with the cognitive processes, tasks can also be placed upon a continuum from 'inducing analytical cognition' to 'inducing intuitive cognition'.

Emotional state can be divided in three levels: valence (pleasure), arousal (energetic) and dominance (control) (Mehrabian, 1996), and is an important factor in decision making (Mosier & Fischer, 2010). Affect can be induced by the decision task itself (integral affect) or can be present beforehand (incidental affect). Incidental affect influences heuristics and the way in which information is processed (Mosier & Fischer, 2010). This type of affect also influences judgments as explained by the Affect Infusion Model by Forgas (1995).

The appraisal and perceived task demand will determine the individual's coping strategy. Research agrees that there are basic coping strategies that will be used when under stress, emotion-focusses and task-focussed coping (Endler & Parker, 1990) and that task-stress triggers different coping styles in different individuals (Matthews & Campbell, 1998). The chosen coping strategy, on its turn, influences how the individual reacts on the situation and
the event and what decisions that will be made by the individual (Delahajj, 2009). Whether the behavioural reaction is appropriate for the stressful event or not will determine the performance on the task.

Although appraisal, decision-making styles and coping strategies fit in the COPE-model, they are out of the scope of this study since there are no quick and easy ways to determine what appraisals fit in the COPE-model, they are out of the scope of this study for the stressful event or not will determine the performance on the event and what decisions that will be made by the individual.

Whether the behavioural reaction is appropriate for the stressful event or not will determine the performance on the event and what decisions that will be made by the individual. It was therefore expected that variation in the two factors could be associated with observed performance variation.

2. Methods

The study had a longitudinal correlational design. Over a period of 520 days, multiple observations were made with regard to emotional state, cognitive task load and task performance on the same set of tasks.

2.1. Participants

A total of six male participants with a mean age of 32.3 (minimum 27, maximum 38 years) were selected for the Mars500 project. The selection procedure required male volunteers between 25 and 50 years of age with a higher education degree. The participants were divided into two groups of three participants. These

executed different tasks. In the COPE-model, ES and subjective CTL (i.e., perceived task demand) influence each other, and determine what the results will be of the task that is being performed. This leads to the following two hypotheses investigated in this study:

1. The cognitive and affective variables in the COPE model are influenced by the work content:
   a. How well a task goal fits a higher level goal, reflects in the levels of the ES and CTL levels.
   b. Different overall mission phases evoke different levels of the ES and CTL levels.

2. Variation in task performance can be explained by individuals’ emotional state and perceived task load.

It was expected that different tasks with different task goals, evoke different levels of ES and CTL. The different task goals can either fit with the higher level goals. If this is the case, valence, arousal and dominance are expected to increase compared to emotional state levels of tasks with unfitting goals. The same can be expected for cognitive task load. If a task is appropriate for reaching a higher-level goal, CTL will increase.

The phase of the mission was also expected to influence these variables since different phases are related to higher-level goals, or not. Phases during a Mars mission are, however, of different nature that mission phases during MIR and ISS missions (Gushin et al., 1993; Manzey, 2004). For one, a Mars mission lasts 520 days instead of 2 or 3 months. Therefore, the influences of mission phases on ES and CTL levels are expected to differ from those in previous studies with MIS and ISS crewmembers. Euphoria caused by returning home, is present in ISS and MIR mission since the return home takes a few days. Returning home from Mars takes approximately 6 months and a euphoric feeling based on a return mission is not expected in the last phase.

- Arousal is expected to decline during the entire mission since crewmembers get adapted to the situation. At the end of the mission, they are not as excited as at the beginning.
- Valence is expected to be quite stable over the mission. Except the period around the Mars landing, where a high valence is expected. Since the mission phases in the Mars500 project are quite long, this effect might not be strong enough to be visible in one phase compared to other phases.
- Dominance is expected to act in the same manner as valence.
- Cognitive task load is expected to decrease over the course of the mission. The tasks have been performed for a while and do not cause as much CTL as before. The perceived task demand might increase at the end of the mission when crewmembers are more fatigued.

According to the COPE model, the cognitive and affective factors influence performances. It was therefore expected that variation in the two factors could be associated with observed performance variation.

1.3. Research questions

In an effort to gather knowledge on psychological effects of a Mars mission on its crewmembers, the European Space Agency (ESA) and the Institute of Biomedical Problems (IBP) carried out the Mars500 project. This project simulated a Mars mission in its full length of 520 days here on Earth including the isolation factors and the lack of contact with Earth. Obviously, the absence of gravity was not reproduced but the unique settings of a Mars mission simulation brought its own unique stress-factors. During the experimental sessions in the Mars500 project, emotional state (ES) and cognitive task load (CTL) were measured while the participants

2. Work content

As mentioned above, the goals that need to be achieved by the task performer are often hierarchical structured. Completing an overall training to learn a certain skill is a higher level goal. Such goals will be accomplished by achieving different lower level goals, or sub-goals, such as learning different components of the skill. For some lower level goals, the link to the higher level goals will not be as obvious as for other lower level goals, e.g., the mapping of the work goals on the computer tasks is not straightforward (cf. Kieras & Polson, 1985; Sutcliffe, Ryan, Doubleday, & Springer, 2000). This mismatch between the different hierarchical goals will be visible in the perception of the work content. In other words, the appraisal of the work content is dependent on the fit between the lower level and higher level goals.

The simulated Mars mission described in this paper also contains goals from different levels. A few higher level goals will be present during different phases of the mission. Long-term missions to the ISS and MIR space stations have been divided according to a stage model by Manzey (2004) and Gushin, Kholin, and Ivanovsky (1993). These missions last between 4 and 6 months and every stage has its own psychological stressors. In the first phase, that last approximately 4–6 weeks, crewmembers mainly focus on adaptation to the physiological changes. Stress and performance problems in this phase are induced by these physical adaptations. Full adaptation to the new conditions is reached in the second phase that is followed by the most difficult third phase, were psychological problems are likely to occur. This third phase starts after approximately 6–12 weeks in space. Severe stressors in this phase are: monotony and (social) boredom, isolation from family and friends, and the omnipresent contact with the other crewmembers. The fourth phase starts shortly before the end of the mission. It evokes euphoria but also concerns as to ending and completing the mission. Within these different stages during a long-term mission, different higher level goals can play a role. In addition to the goal of exploring the Mars surface, the Mars500 mission distinguishes four phases with corresponding (higher level) goals:

1. An adapt towards the (new) space environment, (2) establish efficient work procedures or routines, (3) prepare for the Mars landing, and (4) return to home (Earth). The different computer tasks that need to be performed during the different phase of the missions have lower level goals that, ideally, would contribute to such higher level goals (i.e., support the adaptation, the routine development, the landing preparation and the return).

2. The study had a longitudinal correlational design. Over a period of 520 days, multiple observations were made with regard to emotional state, cognitive task load and task performance on the same set of tasks.

2.1. Participants

A total of six male participants with a mean age of 32.3 (minimum 27, maximum 38 years) were selected for the Mars500 project. The selection procedure required male volunteers between 25 and 50 years of age with a higher education degree. The participants were divided into two groups of three participants. These
were also the groups in which the tasks were performed. For practical reasons, one group consisted of the English speaking participants and one group consisted of the Russian speaking participants.

2.2. Work content

Every other week, a session started for half an hour. In every session three tasks were executed: a learning activity, called Collaborative Trainer (COLT); a negotiation game, named Colored Trails (CT); and an entertainment game, called Lunar Lander (LL). COLT and CT are multi-user (group) tasks, whereas LL is a single-user game. The different tasks are explained in the next sections.

2.2.1. Colored Trails (CT)

Colored Trails is a negotiation game with competitive elements for two or more users. This game is developed as a research test-bed for investigating decision-making in groups and proposed as a tool for assessing group-members’ relationships and (a-) social behaviours towards each other (Gorbunov et al., 2011). The three group members played the game on a rectangular board with coloured squares (see Fig. 2a). Group members had their own piece on the board, which they could move with a coloured chip. The general goal was to position pieces as close as possible to the flag. All players saw the board and the chips possessed by the other players, which made it possible to propose chip exchange. A player who received propositions could either except one or decline all. According to a specific scoring function, each player could earn points with its moves. The game-time was around 10 min (for a more detailed description of the game and analyses of the group-members’ CT-performances and relationships, see Gorbunov et al. (2011)).

2.2.2. Lunar Lander (LL)

Lunar Lander is an entertaining game, played individually. This version was a Java-version of the original 1979’s Lunar Lander video game from Atari. A player had to land a space-ship safely on the surface of the moon as many times as possible without crashing (see Fig. 2b). The difficulty level increased successively. Pressing the arrow buttons altered the space-ships direction and the space-bar accelerated the space-ship forward.

2.2.3. Collaborative Trainer (COLT)

Collaborative Trainer was a learning task for three persons, one teacher (instructor) and two students. The students’ goal was to learn procedures for the usage, maintenance and damage-control of systems. The teacher had to provide the assignments and to guide student’s learning processes (i.e., pointing to the relevant learning material and giving hints when needed). This way, COLT combines computer-based learning and collaborative learning techniques. The teacher sent instructions via chat to the students, who then executed the specific task. For each assignment, the teacher had background information available on his dashboard to supervise, help and advice the students while they were learning. COLT was used to learn the relevant procedures of two different systems: Cardiopres and Watertank. Cardiopres is a real payload for physiological measurements in space stations (ECG, breathing, skin conductance, blood pressure), and COLT contained all “official” procedures and background (multimedia) information for its usage, maintenance and Fault Detection, Isolation and Recovery (FDIR) procedures (see, Fig. 3). The Watertank system was a simplified simulator of a hypothetical water provision system, for which COLT provided some derived procedures for usage and fault recovery. The Watertank scenario was always played with the same teacher, whereas for the Cardiopres the teacher role rotated among the three group members (for more details on COLT, see Smets et al. (2012)).

2.3. Work content: phases

The whole simulated Mars mission lasted for 520 days and was divided into four different phases corresponding the stage model described by Manzey (2004) and Gushin et al. (1993). The simulated Mars landing divided the mission into two halve. Both halves were divided equally, resulting in four phases. The first phase (session 1–9, week 1 to week 18) and the second phase (session 10–19, week 19–38) were before the Mars landing, and phase three (session 20–29, week 39 to week 58) and four (session 30–38, week 59) were after the Mars landing.

2.4. Measures

Several variables were collected to measure the abstract constructs of the COPE model: emotional state, cognitive task load, and task performance.

2.4.1. Emotional state (ES) & cognitive task load (CTL)

A common way of measuring emotional state is by using the Self-Assessment Manikins from Bradley and Lang (1994). This questionnaire consists of three 5-point-likert scales on valence, arousal and dominance. Instructions were given to the participants that explained the scales and the extreme points on the scale. Valence represents pleasure on a scale from negative (e.g., sad) to positive (e.g., happy). Arousal is the participants’ activity level from low (e.g., calm) to high (e.g., excited). The level of dominance represents the amount of control in the situation on a scale from minimal (e.g., guided) to maximal (e.g., autonomous). Every point on the scales was represented by a small icon as shown in Fig. 4.
The three levels of cognitive task load were also measured on a 5-point scale. This questionnaire is shown in Fig. 4 as well. For each level there was a separate question. For Level of Information Processing, the question was how complex is the task for the actor? This could be answered on a scale from low (routine or automatic) to high (non-routine or intensive problem-solving). For Task-Set Switches the question was how often must the actor switch from one task to the other? This could be answered on a scale from low (one task after the other) to high (continuous interleaving or interruption). Time occupied was measured by asking how much of the available time is the actor focused on and occupied with his or her work. This question was answered on a scale from low (slowing down and pauses are allowed) to high (required to perform at maximum speed to complete the tasks).

Next to the rated ES and CTL scores, the difference between two of these scores were also used in the analyses of this study. For example, valence was measured before a task, and after a task. The $\Delta$ valence was used as an indication of valence change. In Section 2.6 the different measurement moments of ES and CTL are explained, and Section 3.1 explains the $\Delta$'s variables in more detail.

2.4.2. Performance measurements

All three tasks aim at achieving individual task or learning goals; the performance scores were determined in different ways. During Lunar Lander, points were received for every successful landing. The score that could be achieved for a landing on a particular spot was visible underneath the surface of that spot as shown in Fig. 3.

For Colored Trails the score was calculated as follows; reaching the goal location would deliver 125 points. For not reaching the goal, 25 penalty points were subtracted for every square between
the goal and the player's position. In addition, for every chip the player had not used, he received 10 extra points.

After performing the COLT tasks, a task questionnaire was filled in, in which the retention of the knowledge gathered during the task was examined. It asks questions about facts and procedures. This was followed by a questionnaire asking the teacher to score the students and himself, and asking him to ask students to rate his teacher performance. Scores were on a 5-point scale: from 1 (poor) to 5 (good). Students also received a similar questionnaire, asking them to rate their own performance, and asking another participant to score their performance.

2.5. Experimental design

The experiment had a repeated measures design. Over a period of 520 days, every other week, multiple observations were made with regard to emotional state, cognitive task load and task performance during the execution of three computer-based tasks.

2.6. Procedure

Every two weeks the groups performed a session for half an hour, consisting of all three games: Lunar Lander, Colored Trails and one of the COLlaborative Trainer tasks. First the participants logged on to the system and a timer, a chat client, and an overview screen started. The timer and chat client were on during the whole experiment. Next, a game performance screen was shown, followed by the first game or evaluation task of that session. All three games followed an almost similar procedure (Fig. 5). The task starts with an emotional state questionnaire (time = T0), followed by the tasks. After the task was completed an emotional state and a cognitive task load questionnaire followed (time = T1). For the Lunar Lander and Colored Trails task, the procedure stops there. The COLT sequence continued with an examination part, followed by a teacher/student questionnaire and a second emotional state and cognitive task load questionnaire (time = T2).

3. Results

All analyses were executed in R Studio and an alpha level of 0.05 was used for all statistical tests. Before the analyses were conducted, the data needed preparation.

3.1. Data preparation

Reliability analyses in Table 1 showed a high level of consistency between the three cognitive task load measures (LIP, TOC, and TSS). The three separate levels were replaced for a single aggregated mean score for cognitive task load that was used as a predictor in the regression analyses and as dependent variables in the ANOVAs. An extra variable was created with the difference between ES measurement at T1 and T0 (valueT1 − valueT0 = valueΔT). Another extra variable was created to indicate the phase of the mission. Phase 1 lasted from session 1 up to and including session 9, phase 2 included sessions 10 up to 19, phase 3 consisted of sessions 20 up to 29 and phase 4 included sessions 30 up to 38.

The small group of participants in this study might be “interesting in themselves” and create a “sample that exhausts the population” which are indications for fixed effects (Gelman, 2005). When this is the case, the participants can be treated as fixed effects (Mirman, Dixon, & Magnuson, 2008). This also avoids the risk of over fitting the data when alternatively a within-subject variable to identify the 38 sessions would have been included. Therefore, in all the ANOVA’s and multiple linear regressions described in the result section, participants were treated as a fixed effect by adding categorial participant variables into the models.

3.2. Task differences

To test the first hypothesis, a series of one-way ANOVA’s were conducted to examine if the cognitive task load and emotional state variables varied when different tasks were executed. For all these ANOVAs tasks was the independent variable with 5 levels, i.e. COLlaborative Trainer (3 versions), Colored Trails, and Lunar Lander. The dependent variables were the aggregate cognitive task load level, and the emotional state levels, i.e. arousal, dominance, and valence, measured at T1 and T2, and ΔT1. The results of the ANOVAs are presented in Table 2 and show a significant effect on cognitive task load at T1, on valence at T2, and on valence at ΔT1 and on arousal and dominance at ΔT2. Additional Tukey’s post-hoc tests showed between which tasks the differences were found. The bar graphs in Fig. 6 show all the significant differences.

3.3. Work content – phase differences

Further investigations into the first hypothesis looked at differences in CTL and ES variables depending on the phase of the mission in which these values were measured. Details of these ANOVA’s are displayed in Table 3. Differences between phases were found for Cognitive Task Load at T1, arousal at T0, T1 and T2, and valence at T0 and ΔT1. For the dominance level of Emotional State no differences between phases were found suggesting that this level did not vary during the simulated Mars mission. Differences

---

**Table 1**

<table>
<thead>
<tr>
<th>Task</th>
<th>Cronbach's alpha mean score</th>
<th>Unstandardized CTL mean score</th>
</tr>
</thead>
<tbody>
<tr>
<td>COLT T1</td>
<td>0.92</td>
<td>0.94</td>
</tr>
<tr>
<td>COLT T2</td>
<td>0.96</td>
<td>0.95</td>
</tr>
<tr>
<td>Lunar Lander T1</td>
<td>0.93</td>
<td>0.98</td>
</tr>
<tr>
<td>Colored Trails T1</td>
<td>0.87</td>
<td>0.93</td>
</tr>
</tbody>
</table>

**Fig. 5.** Procedure for the different tasks. Top picture shows lunar lander and colored trails procedure, bottom picture shows the procedure for the COLT tasks.
Table 2
Results of the ANOVAs showing effects of task on CTL, arousal, dominance and valence at different measurement moments.

<table>
<thead>
<tr>
<th></th>
<th>df1</th>
<th>df2</th>
<th>Sum of squares</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTL1</td>
<td>4</td>
<td>512</td>
<td>11.68</td>
<td>6.93</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>CTL2</td>
<td>2</td>
<td>136</td>
<td>2.23</td>
<td>2.59</td>
<td>0.079</td>
</tr>
<tr>
<td>T0 arousal</td>
<td>4</td>
<td>515</td>
<td>3.19</td>
<td>2.05</td>
<td>0.086</td>
</tr>
<tr>
<td>T0 dominance</td>
<td>4</td>
<td>515</td>
<td>0.38</td>
<td>0.19</td>
<td>0.943</td>
</tr>
<tr>
<td>T0 valence</td>
<td>4</td>
<td>515</td>
<td>0.30</td>
<td>0.54</td>
<td>0.706</td>
</tr>
<tr>
<td>T1 arousal</td>
<td>4</td>
<td>513</td>
<td>0.49</td>
<td>0.43</td>
<td>0.787</td>
</tr>
<tr>
<td>T1 dominance</td>
<td>4</td>
<td>513</td>
<td>1.51</td>
<td>1.64</td>
<td>0.162</td>
</tr>
<tr>
<td>T1 valence</td>
<td>4</td>
<td>513</td>
<td>3.03</td>
<td>1.54</td>
<td>0.190</td>
</tr>
<tr>
<td>ΔT1 arousal</td>
<td>4</td>
<td>516</td>
<td>0.32</td>
<td>0.36</td>
<td>0.837</td>
</tr>
<tr>
<td>ΔT1 dominance</td>
<td>4</td>
<td>516</td>
<td>1.42</td>
<td>2.32</td>
<td>0.056</td>
</tr>
<tr>
<td>ΔT1 valence</td>
<td>4</td>
<td>516</td>
<td>3.13</td>
<td>3.14</td>
<td>&lt;.014</td>
</tr>
<tr>
<td>T2 arousal</td>
<td>2</td>
<td>136</td>
<td>0.002</td>
<td>0.003</td>
<td>0.997</td>
</tr>
<tr>
<td>T2 dominance</td>
<td>2</td>
<td>136</td>
<td>0.69</td>
<td>1.49</td>
<td>0.229</td>
</tr>
<tr>
<td>T2 valence</td>
<td>2</td>
<td>136</td>
<td>3.88</td>
<td>4.23</td>
<td>&lt;.017</td>
</tr>
<tr>
<td>ΔT2 arousal</td>
<td>2</td>
<td>145</td>
<td>66.42</td>
<td>5.42</td>
<td>&lt;.005</td>
</tr>
<tr>
<td>ΔT2 dominance</td>
<td>2</td>
<td>145</td>
<td>71.94</td>
<td>3.35</td>
<td>&lt;.038</td>
</tr>
<tr>
<td>ΔT2 valence</td>
<td>2</td>
<td>145</td>
<td>133.30</td>
<td>2.38</td>
<td>0.097</td>
</tr>
</tbody>
</table>

**p < 0.001, *p < 0.01, *p < 0.05, p < 0.1.

at T0 are most interesting, as the ES levels have not yet been affected by executing tasks.

Tukey's posthoc tests were conducted to examine difference between phases in more detail. Three expected increases or decreases were found. Arousal decreases from phase 2 to phase 3 at T0 and T2 (Fig. 7a and e). Valence at T0 increased between phase 1 and 2 (Fig. 7b). More differences were found between the non-adjacent phases shown in the bar graphs presented in Fig. 7.

3.4. Explaining task performance variation

To find predictors for performance score, regression analyses were conducted. The emotional state and cognitive task load variables showed differences between tasks, therefore, different regression analyses were conducted per task. Table 4 shows the results of regression analyses for Lunar Lander task and Colored Trails task. The model for Lunar Lander was able to account for 40.5% of the performance variance, \( F(12, 175) = 11.59, p < 0.001 \). The dominance levels before the task, and the change in dominance were significant predictors and both had a positive association with the
Table 3
Results of the ANOVAs showing effects of the different phases on CTL, arousal, dominance and valence at different measurement moments.

<table>
<thead>
<tr>
<th>Phase</th>
<th>df1</th>
<th>df2</th>
<th>Sum of squares</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTL1</td>
<td>3</td>
<td>513</td>
<td>3.99</td>
<td>3.06</td>
<td>*0.03</td>
</tr>
<tr>
<td>CTL2</td>
<td>3</td>
<td>135</td>
<td>0.72</td>
<td>0.54</td>
<td>0.65</td>
</tr>
<tr>
<td>Arousal</td>
<td>3</td>
<td>516</td>
<td>4.84</td>
<td>7.31</td>
<td>***&lt;0.001</td>
</tr>
<tr>
<td>Valence</td>
<td>3</td>
<td>516</td>
<td>0.71</td>
<td>1.76</td>
<td>0.16</td>
</tr>
<tr>
<td>T0</td>
<td>3</td>
<td>516</td>
<td>4.59</td>
<td>3.58</td>
<td>**0.01</td>
</tr>
<tr>
<td>Arousal</td>
<td>3</td>
<td>514</td>
<td>4.87</td>
<td>5.88</td>
<td>***&lt;0.001</td>
</tr>
<tr>
<td>Valence</td>
<td>3</td>
<td>514</td>
<td>1.31</td>
<td>1.91</td>
<td>0.13</td>
</tr>
<tr>
<td>T1</td>
<td>3</td>
<td>514</td>
<td>1.89</td>
<td>1.23</td>
<td>0.28</td>
</tr>
<tr>
<td>Arousal</td>
<td>3</td>
<td>517</td>
<td>0.11</td>
<td>0.17</td>
<td>0.92</td>
</tr>
<tr>
<td>Valence</td>
<td>3</td>
<td>517</td>
<td>0.36</td>
<td>0.78</td>
<td>0.51</td>
</tr>
<tr>
<td>ΔT1</td>
<td>3</td>
<td>517</td>
<td>2.22</td>
<td>2.96</td>
<td>*0.03</td>
</tr>
<tr>
<td>Arousal</td>
<td>3</td>
<td>135</td>
<td>3.48</td>
<td>4.49</td>
<td>***&lt;0.005</td>
</tr>
<tr>
<td>Valence</td>
<td>3</td>
<td>135</td>
<td>0.86</td>
<td>1.24</td>
<td>0.30</td>
</tr>
<tr>
<td>T2</td>
<td>3</td>
<td>135</td>
<td>2.12</td>
<td>1.49</td>
<td>0.22</td>
</tr>
<tr>
<td>Arousal</td>
<td>3</td>
<td>144</td>
<td>0.57</td>
<td>0.39</td>
<td>0.76</td>
</tr>
<tr>
<td>Valence</td>
<td>3</td>
<td>144</td>
<td>0.60</td>
<td>0.39</td>
<td>0.76</td>
</tr>
<tr>
<td>ΔT2</td>
<td>3</td>
<td>144</td>
<td>1.85</td>
<td>0.65</td>
<td>0.58</td>
</tr>
</tbody>
</table>

*p < 0.001, **p < 0.01, *p < 0.05, p < 0.1.

Variation in task performance. No significant model was found for the Colored Trails task, \( F(12, 172) = 0.99, p = 0.46 \).

The three COLT versions resulted in the three models shown in Table 5. For these tasks, an additional difference between T2 and T1 (\( \text{value}_{T2} - \text{value}_{T1} = \text{value}_{\Delta T2} \)) was entered as an indicator for emotional change associated with the examination part. None of the variance in task performance during the Cardiopres task could be accounted for by the variables in the model (\( F(13, 15) = 1.33, p = 0.30 \)).

For the teacher task, 64.4% of the variance in task performance could be attributed to the model (\( F(16, 29) = 6.08, p < 0.001 \). The model intercept and cognitive task load at T2 had a positive relation with task performance for this task. CTL at T1 also showed a trend.

Variance in task performance during the Watertank task could be accounted for by the variable in the model for 74.8% (\( F(15, 50) = 13.87, p < 0.001 \)). The model intercept, arousal at T0 and both valence at \( \Delta T1 \) and \( \Delta T2 \) had a positive influences on task.
4. Discussion

First of all, it should be noted that the results and conclusions are based upon a small number of participants. The participants, however, went through a selection procedure that made sure that they fitted the profile of crewmembers that would actually go on a Mars mission. The Mars500 project is a unique experimental environment and it would be interesting to see if similar projects with different participants would confirm the our findings.

The first hypothesis states that work content influences the cognitive and affective factors. This hypothesis was divided over two aspects of work content. The higher level goals (different mission phases), and lower level goals (different tasks). The findings provided support for these hypotheses since the CTL and ES values varied significantly between different tasks and different phases of the Mars simulation.

It was expected that tasks of which the task goals had a good fit with the higher level goals, caused higher emotional state and cognitive task load levels. Following these expectations, the COLT tasks are expected to have higher CTL and ES values since these tasks have goals fitting in the Mars mission environment. These expectations are found in the higher CTL levels for the teacher task compared to the Lunar Lander and Colored trails tasks. The Watertank task was not a realistic task and might therefore score lower compared to the Lunar Lander and Colored trails tasks. The Water tank task was not a realistic task and might therefore score lower compared to the Lunar Lander and Colored trails tasks. The Water tank task was not a realistic task and might therefore score lower compared to the Lunar Lander and Colored trails tasks. The Water tank task was not a realistic task and might therefore score lower compared to the Lunar Lander and Colored trails tasks. The Water tank task was not a realistic task and might therefore score lower compared to the Lunar Lander and Colored trails tasks.

4.4. Teachers

The first hypothesis states that work content influences the cognitive and affective factors. This hypothesis was divided over two aspects of work content. The higher level goals (different mission phases), and lower level goals (different tasks). The findings provided support for these hypotheses since the CTL and ES values varied significantly between different tasks and different phases of the Mars simulation.

It was expected that tasks of which the task goals had a good fit with the higher level goals, caused higher emotional state and cognitive task load levels. Following these expectations, the COLT tasks are expected to have higher CTL and ES values since these tasks have goals fitting in the Mars mission environment. These expectations are found in the higher CTL levels for the teacher task compared to the Lunar Lander and Colored trails tasks. The Water tank task was not a realistic task and might therefore score lower on several valence measures compared to Lunar Lander and Cardiopresses. The Cardiopresses task on the other hand, had lower arousal and dominance than the Watertank task.

During different phases of the mission, cognitive and affective states had different values. At three instances the directions of these changes were as predicted. For example, the findings confirm that arousal decreased after the landing, suggesting adaptation or maybe boredom, between phases two and three (at T0 and T2). Next, after the initial adaptation and the prospect of landing on Mars, valence went up between phase one and two (at T0). Besides the findings for phase effects, the results revealed several other differences over non-adjacent phases. Important to note is that none of these findings contradicted hypothesised directions. For example, the perceived level of cognitive task load (T1) was higher in the fourth phase compared to the second phase suggesting that tasks are seen as more difficult or more demanding during the last months of this simulated mission.

The second hypothesis stated that the cognitive and affective factors are predictors for task performance. The strongest support for this hypothesis was provided by the findings of the Watertank task, which showed valence as a significant predictor after Bonferroni corrections were applied to the results. These findings support the thought explained by Matthews, Davies, Westerman, and Stammers (2008), that stress reactions are distinguished on the characteristics of the challenge that is faced. Task characteristics determine the task goals and the fit with higher-level goals determine appraisal of the task and emotional states. Matthews et al. (2008) investigated the concept of Lazarus and Folkman and found that changes in stress state induced by tasks, varied with task demands. Matthews et al. (2008) also state that subjective responses are influenced by a person’s appraisal of the task and the environmental demands. When a task is appraised as overloading it will evoke stress, but when a task is appraised as a challenge it will evoke task engagement.
While the COPE model looks at a person's appraisal when faced with a stressful situation, one's motivation to solve a problem or execute a task is left out. However, literature explains that optimal performance on a task is also related to the amount of interest someone has for performing that task (O'Keefe & Linnenbrink-Garcia, 2014). Particularly the differences between the Cardiopres and Watertank exercises in the COLT-task can be explained by differences in experienced challenge and interests. Whereas the Cardiopres is a real (existing) payload for manned space missions, the Watertank was a simplified simulation of a complex system which the crew-members did not experience as realistic (e.g., leading to a different valence–arousal relationship).

Effects of work content on cognitive and affective processes and their effect on performances were found, even though the environment contained noise. The COPE model seems therefore a good starting point to dynamically adjust the work content to (predicted) cognitive, affective or performance factors. In the space domain, we are improving ePartner's (objective) real-time monitoring functions of crewmember's cognitive task load and emotional state, which is expected to provide better predictions for the task performance. In the space and naval domain, COPE-based support functions, like real-time feedback on the emotional state and error risks, are being designed and tested.

To appreciate the results of this study, a number of limitations should be considered. (1) Only six participants performed the tasks and they were selected on specific qualities as described in the Method Section. A previous study on the same dataset revealed that there might be cultural differences between the two groups of three participants (Smet et al., 2012). They did, however, participate in many sessions and thus provided a large set of data. To overcome any of these biases, all analyses accounted for individual user variation (Mirm et al., 2008). (2) The emotional state and cognitive task load measures were all subjectively measured. The results can therefore contain social desirable answers and noise due to momentary perception biases. Since the COLT task asked for the teacher to rate the student and vice versa, the results of the COLT tasks might be biased. Mark A. Neerincx, Kennedie, Grootjen, and Grootjen (2009) created predictive models with objective cognitive task load predictors and reached a higher accuracy in those predictions for the crewmembers of a naval Ship Control Centre. (i.e., from 74% to 86% accuracy). Not only did they not experience social desirable answers, the data was also collected during the task and not afterwards. (3) The COPE model looks at the current state of a person when engaged in a stressful task. However, in this study participants were asked to rate their state in the task afterwards. (4) Although the findings did show significant regression models to explain task performance, for one of the collaborative training task, i.e. Cardiopres, and the negotiation and collaboration learning game, i.e. Colored Trails, the COPE factors did not lead to a significant regression model.

5. Conclusion

In the present paper, data from the Mars500 project was fitted to the COPE-model to investigate influences on cognitive and affective measures, and those factors' predictability on task performance. The Mars500 project, a simulated Mars mission that lasted for 520 days, inflicts unique stressors such as social isolation, incidents and boredom on its crewmembers. Different tasks were performed during this mission, while gathering subjective emotional state and cognitive task load data.

The findings support the general hypothesis that work content, i.e. different tasks and mission phases, can influence cognitive and affective factors, and that these factors, on their turn, can explain task performance. Designing work for, or adjusting work plans during, long-term missions could benefit from this insight by considering cognitive task load and emotional state when rescheduling tasks. The findings also give some insight in the validity of the COPE model, showing the relation between external work factors, internal cognitive and cognitive factors and eventually the external performance when operating under stress.

Acknowledgements

This research (project number 056-22-010) is supported by the Dutch FES program “Brain and Cognition: Societal Innovation”. Part of this research was conducted for the MEGA project. The MEGA research and development program is funded by the European Space Agency (contract numbers 19149/05/NL/JA and 21947/08/NL/ST).

References


