The Mission Execution Crew Assistant: Improving Human-Machine Team Resilience for Long Duration Missions

M.A. Neerincx, J. Lindenberg, N. Smets
TNO /DUT, The Netherlands
mark.neerincx@tno.nl

A. Bos, L. Breebaart, T. Grant
S&T BV /NDA
The Netherlands
bos@stcorp.nl

A. Olmedo- Soler
OKS
Spain
aos@olmedo.com

U. Brauer
EADS-ST
Germany
Uwe.Brauer@astrium.eads.net

M. Wolff
ESA-ESTEC
The Netherlands
mikael.wolff@esa.int

ABSTRACT

Manned long-duration missions to the Moon and Mars set high operational, human factors and technical demands for a distributed support system, which enhances human-machine teams' capabilities to cope autonomously with unexpected, complex and potentially hazardous situations. Based on a situated Cognitive Engineering (sCE) method, we specified a theoretical and empirical founded Requirements Baseline (RB) for such a system (called Mission Execution Crew Assistant; MECA), and its rational consisting of scenarios and use cases, user experience claims, and core support functions. The MECA system comprises distributed personal ePartners that help the team to assess the situation, to determine a suitable course of actions to solve a problem, and to safeguard the astronauts from failures. In addition to standard requirements reviews, we tested and refined the RB via storyboarding and human-in-the-loop evaluations of a simulation-based prototype in a virtual environment with 15 participants. The evaluation results confirmed the claims on effectiveness, efficiency, satisfaction, learnability, situation awareness, trust and emotion. Issues for improvement and further research were identified and prioritized (e.g., acceptance of mental load and emotion sensing). In general, the sCE method provided a reviewed set of 167 high-level requirements that explicitly refers to the tested scenarios, claims and core support functions on health management, diagnosis, prognosis & prediction, collaboration, resource management, planning, and sense-making. A first version of an ontology for this support was implemented in the prototype, which will be used for further ePartner development.

1. INTRODUCTION

Future manned missions to the Moon or Mars take place in hostile and complex environments, have a long duration and may show long delays in communication with the Earth. These missions set high requirements for crew operation support, concerning (a) the scheduling and execution of crew operations to maintain acceptable levels of effectiveness, efficiency and safety, and (b) the accommodation of user characteristics, tasks and contexts to provide the "right" information, services and functions at the "right" time and in the "right" way. For such missions, support systems have to be developed that enhance crew safety and performance during all operations. The Mission Execution Crew Assistant (MECA) project aims at such a system by empowering the cognitive capacities of human-machine teams during planetary exploration missions in order to cope autonomously with unexpected, complex and potentially hazardous situations. An elaborate and sound requirements analysis has been conducted, focusing on a manned mission to Mars. It should be noted that the project outcomes are of relevance for manned space missions where a greater need for autonomy exists (i.e., most outcomes also apply to Moon missions, and a substantial part is relevant for International Space Station missions and ground-based control missions of planetary robots).

1.1 Human-Machine Collaboration

Technological developments are causing a fundamental change of machine's role in complex work environments. They are becoming part of cognitive systems that consist of human and synthetic actors who collaborate for successful attainment of their joint operation objectives (e.g., [1]). To enhance the Human-Machine Collaboration, MECA will act as electronic partner, helping the crew to assess the situation, to determine a suitable course of actions to solve a problem, and to safeguard the astronaut from failures [2]. This concept comprises a collection of distributed and connected personal ePartners to support the (distributed) crew members during exploration missions. A personal ePartner predicts its crew members momentary support needs by on-line gathering and modelling of human, machine, task and context information. Based on these models, it attunes the user interface to these needs in order to establish optimal human-machine performance. The user interface of the ePartner is "natural or intuitive" by expressing and interpreting communicative acts based on a common reference of the human and machine actors.

1.2 Situated Cognitive Engineering

Neerincx & Lindenberg [3] developed a situated Cognitive Engineering (sCE) method for HMC design. Corresponding to the "classical" CE methods [4, 5, 6], it consists of an iterative process of generation, evaluation and refinement. In addition, the sCE method combined the classical human-centered perspective with a technology-centered perspective to address the adaptive nature of both human and synthetic actors with their reciprocal dependencies systematically [7]. Furthermore, the sCE method includes an explicit transfer and refinement of general state-of-the-art theories and models—which include accepted features of human cognitive, affective and social processes—into situated support functions for the specific operational contexts (cf. [8]). Application of the sCE method results in a sound design knowledge base—possibly including design patterns, software frameworks and algorithms for core support functions—with corresponding best practices for the application domain. The Mission Execution Crew Assistant (MECA) project applied this method in order to establish a theoretically sound and empirically proven Requirements Baseline (RB). The process of specification, refinement and validation is based on three information or feedback sources (Figure 1):

1. A work domain and support analysis to identify operational, human factors and technological challenges of manned planetary space missions and, subsequently, to derive the general support concept out of it (see section 2).
2. Expert and task-analytical reviews to assess the RB itself and its rational (i.e., scenarios, claims and core functions; see section 3).
3. Scenario-based prototype evaluation of MECA's claims and core functions with simulation-based prototype (see section 4).

It is important to note that this study focuses on a future MECA system that is in a very early development stage, in which human factors aspects, operational demands and technology are systematically explored. This study follows an iterative human-centered development process corresponding to recent human-factors engineering methods and standards (e.g., ISO 13407 "Human-centered design processes for interactive systems", and the ECSS-E-ST-10-11C "Space Engineering: Human Factors Engineering" standard of the European Cooperation for Space Standardization (ECSS)). Following these methods, the MECA requirements are assessed and refined from rather high-level specifications in early development stages to detailed definitions in late development stages. Consequently, the present requirements baseline contains relatively abstract specifications that will be assessed and refined further.

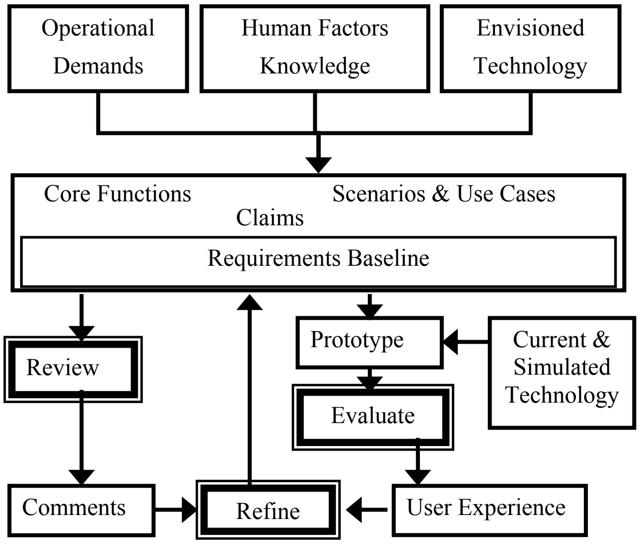


Figure 1: The iterative process of requirements analysis.

2. WORK DOMAIN AND SUPPORT ANALYSIS

2.1 Operations

We conducted a mission and domain analysis, consisting of (1) a mission analysis of previous and future operations of manned planetary space operations, and (2) a work domain analysis to construct an Abstraction Decomposition Space (ADS). For the mission analysis, we assessed documents that contain "lessons learned" from previous missions (e.g., [9], [10]) and studies on future mission operations (e.g., [11-15]). Although astronauts have excellent competencies, are well-trained and have a high work motivation, their performance can deteriorate due to diminished motor, perceptual and cognitive capacities, and emotional or social-psychological problems. There is a clear need to support the crew for both nominal and off-nominal operations, among other things to accommodate team-members creative problem solving processes and counter-balancing initiatives to share or take over specific critical tasks.

Independent from the mission analysis, a Work Domain Analysis (WDA) was performed [16]. WDA is the first phase of Cognitive Work Analysis, a method that aims at socio-technical systems that can fully exploit the problem-solving capabilities and adaptability of human resources in unexpected situations, improving efficiency, productivity and safety. The analysis is event-independent, based on the notion that it is impossible to predict all possible system states. Systems are therefore defined in terms of their environmental and cognitive constraints: their physical environment, priorities, and functionality [17], [18]. For MECA, the WDA provided an Abstraction Decomposition Space offering substantial insight into the properties or, and relationships within the Mars surface mission system (ADS; [16]). Outcomes were compared with outcomes of the mission analysis and the first version of the Requirements Baseline (RB). The main conclusion was that the RB considered all technical aspects of the Mars surface mission in sufficient detail, but that further specifications were required in the areas of general living.

2.2 Human Factors

Parallel to the operations analyses, a literature study was conducted on Human Factors of complex high-demanding task environments, in which well-trained human operators may act in extreme and hostile situations (such as the defense and safety domains). This study provided key issues that MECA should address for supporting the Human-Machine Collaboration: cognitive task load [8], situation awareness [19], sense-making [20], decision making [21], diversity of cognitive capacities [22], trust [23], emotion [24], collaboration [25], and crew resource management [26]):

Cognitive task load. Due to required combination and possibly complexity of different supervision and control tasks, the momentary mental load of the astronauts can be suboptimal. MECA should support an adequate load scheduling over time and available human-machine resources based on a model of cognitive task load that distinguishes three load factors: percentage time occupied, number of task switches and task complexity.

Situation awareness and sense making. If severe, unexpected and time-critical problems occur, time may be limited to communicate with experts at the earth-based Mission Control Center (MCC). In these cases the astronaut and MECA will have to be self-sufficient as a team in their problem-solving capabilities. MECA will have to provide situation awareness and sense making support by helping the astronauts in collecting relevant data, interpreting the data and the generation and testing of hypotheses.

Decision making. Rational decision making shall be supported by exhaustive evaluation of options, collecting and providing an overview, ranking the options, and possibly proposing the best. Naturalistic decision-making is supported by functions that assess the situation based on patterns capturing experience and preference of the crew, recommend actions based on the patterns, check that the execution of the course of actions is according to expectancies, and test assumptions underlying human naturalistic decision-making.

Diversity of cognitive capacities. The astronauts have different expertise and experiences, and will perform their various tasks in different environments. This causes differences in the momentary capabilities, levels of attention and available modalities. MECA needs to be aware of these factors to be able to support the astronaut effectively by tailoring the communication to the available attentional resources and modalities.

Trust and emotion. Given the dependency of the astronauts on MECA and the ways the human-machine collaboration will be shaped, a high level of trust is required. Such a level of trust can be realized with explanatory user interfaces that provide information of the inner workings of MECA. Furthermore, MECA should take into account the user’s emotional state and possible effects of the human-computer interaction on these states.

Collaboration. MECA shall help to avoid gaps and overlap in individuals' assigned work (i.e. support coordination), to obtain mutual benefits of human and machine actors by sharing or partitioning work (i.e. support cooperation) and to achieve collective results that the participants would be incapable of working alone. Furthermore, it shall help to support the generation and maintenance of a shared mental model within human-machine teams, which contains both team knowledge as well as situation knowledge. By mediating between actors, insight will be provided into the other actors’ goals, intentions behavior and needs.

Crew resource management (CRM) is a combination of techniques used to train a crew to operate in complex environments where teams interact with technology, aiming to minimize the effects of errors related with human factors (including communication and cultural aspects) and to maximize the crew effectiveness. MECA should manage the skills and task performance of the crew, and plan and support training to keep performance to a level required by operational demands.

2.3 Technology

In addition to the operations and human-factors analysis, we conducted a technology assessment that distinguished three types of development that will influence future Human-Machine Collaboration (HMC).

1. Developments in the areas of ubiquitous computing, ambient intelligence and mixed reality bring forth *distributed smart task environments* with completely new ways of (computer mediated) human-human and human-technology interaction (e.g., [27, 28, 29]). Such environments support (a) the access to, and sharing of information, services and resources, and (b) the presence and engagement in distributed team activities.

2. Current cross-disciplinary Research and Development (R&D) communities bring forth *enabling technology* that can be integrated, and possibly embedded, in such environments in order to sense, interpret and anticipate individual human conditions and behaviors (e.g. to improve safety and health) around themes like pervasive, affective and mobile computing; persuasive technology; augmented cognition; networked tangible interfaces ("Internet of Things"); and human-robot interaction (e.g., [30-35]).

3. Developments in Artificial Intelligence, such as Multi-Agent Systems, bring forth real or virtual machines that can act autonomously in dynamic environments and take the initiative in *joint human-machine operations* (e.g., [36-38]).

There is one essential assumption for MECA derived from this assessment, i.e., it will act in a Smart Task Environment with automatic distribution of data, knowledge, software and reference documents. However, MECA should still provide operational support—based on history and current available information and knowledge—when infrastructure failures occur. It will apply state-of-the-art Agent and Web technology, model-based reasoning and health management, human-machine (e.g. robot) collaboration and mixed reality. An interesting new approach that fits very well with the MECA framework is the concept of self-assembling wireless autonomous reconfigurable modules [39]. Boundaries for this situated technological design space are set by the technical requirements on maturity, graceful degradation, maintainability and fault tolerance, and the fact that the technology will be performing in a hostile environment. Shared ontologies are seen as an enabling infrastructure for MECA to effectively enhance human-machine collaboration.

3. SITUATED REQUIREMENTS BASELINE

Section 2 summarized the first component of the sCE methodology (i.e., a work domain and support analysis). The second component of this methodology consists of the construction and maintenance of the Requirements Baseline (a table with all requirements) and the general design rationale that consists of the core functions, claims, and scenarios & use cases.

3.1 Design rationale

For the specification of the design rationale, we distinguish three steps that should be followed both from top and bottom (i.e., a top-down, goal-directed approach to work out core functions and a bottom-up, event-driven approach to address contextual or situational demands in scenarios, [8]).

First, the *core functions* of the system are derived from the WDS analysis. Based on the analyses of operational demands and human-factors aspects, we identified seven 'core functions' that MECA should accomplish with the envisioned enabling technology. These core functions are: health management, diagnosis, prognosis & prediction, collaboration, resource management, planning, and sense-making

Second, for each core function, one or more testable *claims* on its operational effects have to be specified; such a claim can be assessed unambiguously in review or evaluation processes. Both positive and negative claims can be specified (cf. [40, 41]). The first three claims consist of standard usability measures, while the subsequent three measures concern additional human experience and knowledge measures:

- Effectiveness and efficiency will be improved both for nominal situations and anomalies, because MECA extends astronauts cognitive resources and knowledge (e.g. procedure and planning).
- Astronauts will express high satisfaction for the MECA support, because (a) it is based on human-machine partnership principles (e.g., for sharing of knowledge) and (b) the astronauts remain in control.
- Working with MECA will be easy to learn, because the support is integrated into the task execution and can be accessed via intuitive multimodal user interfaces (application of visual, auditory and tactile modalities).
- Situation awareness will be enhanced by an overview of relevant situation knowledge with the current plan, and the provision of context-sensitive notification mechanisms.
- Astronauts show appropriate trust levels for MECA, because they share knowledge via simple "ecological" models that astronaut can easily access and understand.
- MECA accommodates emotional responses appropriately in critical situations.

Third, *scenarios & use cases* have to be specified. Scenarios are coherent and situated stories about how a specific team of human and syntehtic actors behave or will behave in specific circumstances with the operational consequences. Scenarios are often visualized via annotated storyboards to present and discuss the design rationale, and used for the evaluation of prototypes. Use cases describe the general behavioral requirements for software systems or business processes, and have a specific specification format. According to our methodology, each use case should explicitly refer both to one or more requirements and to one or more claims. In addition, each claim and each requirement should be included in one or more use cases. A scenario can be viewed as an instance of one or more use cases. Use cases and scenarios are very useful when discussing a not-yet-existing system with different stakeholders. With minor help most people are able to understand these design specifications. For MECA, we specified a rich and diverse set of scenarios. Subsequently, we selected a "core" scenario out of the scenario set, which incorporated all seven 'core functions', and worked this scenario out in more detail. Figure 2 presents an overview of the teams and actors in the MECA core scenario: Team A and B perform extravehicular activities on the Mars surface, while Herman is in the Habitat to perform some maintenance and self-care tasks (rover 2 is moving from Team A to Team B during the scenario).

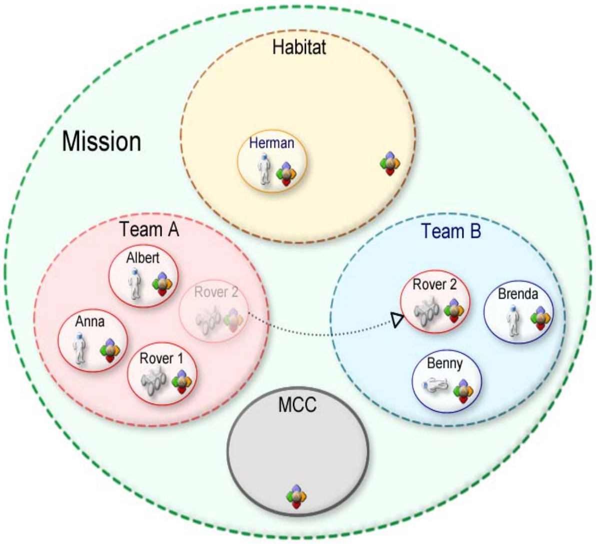


Figure 2: Teams and actors in the MECA scenario. The icon ◆ represents a MECA Unit.

3.2 Annotated requirements

The 'core function', claims and scenario generation support the elicitation, validation and refinement of requirements. For each core function, claims and scenario, one or more requirements have to be specified for the future system (i.e., what the system must do in specific contexts with the expected consequences). The complete set of requirements comprises the requirements baseline. Below is a brief example of a requirement that is linked to several use cases (UC078, UC091 and UC093).

RG1015	MECA shall enhance crew capabilities to perceive their environment, including equipment (health) state and crew (health) state, at a level of detail required for guaranteeing the safety of crew and the success of the mission. *Use Cases: UC078, UC091, UC093*

Table 1 provides a brief overview of the so-called functional requirements. The MECA 'core' scenario incorporates all core functions mentioned in section 3.1, and includes all processes and corresponding functional requirements of Table 1.

Table 1: Outline of functional requirements.

Process	**Functional requirement**
Information Gathering	detect needs for operations and training
Goal Setting	select and prioritize goals for operations and training
Plan Generation or Selection	generate plans, or select pre-generated plans and procedures, for operations and training
Plan Evaluation	evaluate operational and training plans
Prepare for Execution	prepare the resources for executing operational and training plans
Execution	execute operational and training plans
Processing Evaluation of Results	evaluate execution results for operational and/or training purposes

The use cases contain a link to the requirements it satisfies. Table 2 contains a simplified use case for the MECA 'core' scenario that links to the requirement RG1015.

Table 2: Simplified use case description with explicit links to the requirements.

<table>
<tr><th>UC078</th><th>Hypothermic astronaut</th></tr>
<tr><td>Goal</td><td>Treat hypothermic astronaut that is on EVA, arrange that doctor is in medical facility to treat astronaut</td></tr>
<tr><td>Actor</td><td>personal MECA, MECA habitat, Rover, Astronaut in habitat, Astronauts on EVA, hypothermic astronaut</td></tr>
<tr><td>Level</td><td>Level 0 (High Level Ops Scenarios)</td></tr>
<tr><td>Precondition</td><td>Astronaut on EVA is hypothermic</td></tr>
<tr><td>Post condition</td><td>Hypothermic astronaut is in medical facility being treated</td></tr>
<tr><td>Main Success Scenario</td><td><table><tr><th>Step</th><th>Action</th></tr><tr><td>1
…</td><td>personal MECA detects hypothermia in astronaut on EVA …</td></tr></table></td></tr>
<tr><td>Satisfies Requirements</td><td>RG1015, RF2097, RF2110, RF2120, RF2130, RF2131, RF2133, RF2140, RF2163…</td></tr>
</table>

Table 3: Outline of the core scenario with corresponding claims.

Scenario	Claims
1. Benny and Brenda are rock collecting …	
2. Benny's spacesuit heater fails.	Emotion ↓ Situation awareness ↑
3. Benny and MECA unit detect and diagnose problem together.	Effectiveness ↑ Efficiency ↑ Easy-to-learn ↑
4. Brenda helps Benny by looking if there is ice on his suit.	Emotion ↓ Situation awareness ↑
5. MECA predicts hypothermia and calls for help ….	Situation awareness ↑ Satisfaction ↑
6. MECA advises Team B to start walking ….	Trust ↑
7. Herman and MECA select rover to pick up Team B	Situation awareness ↑
8. Herman in the habitat prepares to receive the astronaut.	Easy-to-learn ↑
9. MECA informs astronaut (and other entities) of plan.	Situation awareness ↑
10. Benny faints earlier than predicted.	Trust ↓
11. Brenda, MECA and rover devise way to pick up Benny.	Learn., emotion ↑ Effectiveness ↑ Efficiency ↑
12. Rover transports ….	Situation awareness ↑

Table 3 presents an overview of the annotations of the user experience claims on the effects of MECA support for the core scenario.

3.3 Expert reviews

A System Requirements Review (SRR) was conducted with eight experts from ESA-ESTEC. The experts evaluated the requirements baseline, use cases and scenarios on consistency, understandability, relevance, unambiguousness, verifiability, completeness and coherence. Via a website the experts filled in an individual review "table". The review outcomes were discussed in a group setting to determine which improvement actions should be taken: addition, deletion or change of requirements, and refinement of use cases. Requirements that were added have a trace to the specific input of a reviewer in the table. See an example of such a trace below, with left the unique requirement number, then the requirement and below the trace to the SRR:

RF2021	MECA shall support information gathering to detect needs for operations and training. *Trace: RID:SRR#76; RID:SRR#77*

The start of the second type of cycle "scenario-based prototype evaluation" (see figure 1) consisted of reviewing the design of the MECA system in a Preliminary Design Review (PDR), with the "core scenario" and corresponding claims. This review was conducted by four experts of ESA-ESTEC. The procedure of reviewing was the same as with the SRR, filling in a review table and participating in a meeting to determine the corresponding actions to be taken. This resulted in refinement of the design and scenario.

4. PROTOTYPE EVALUATIONS

Based on the analysis and review outcomes of section 3, we consolidated the scenario and developed a general architecture that is the starting point for prototype (see Figure 3). The core functions of MECA will be incrementally included in this prototype. For the scenario and support functions, we constructed a storyboard. This storyboard is the vehicle to realize our incremental incorporation of MECA functions into the prototype: more-and-more the functions—as visualized in the storyboard—replace the corresponding storyboard components. The first incomplete prototype was complemented with a storyboard for a major part to allow for a first evaluation via a user group walkthrough, whereas the last version only included the "completion storyboard picture" to allow for simulation-based evaluation.

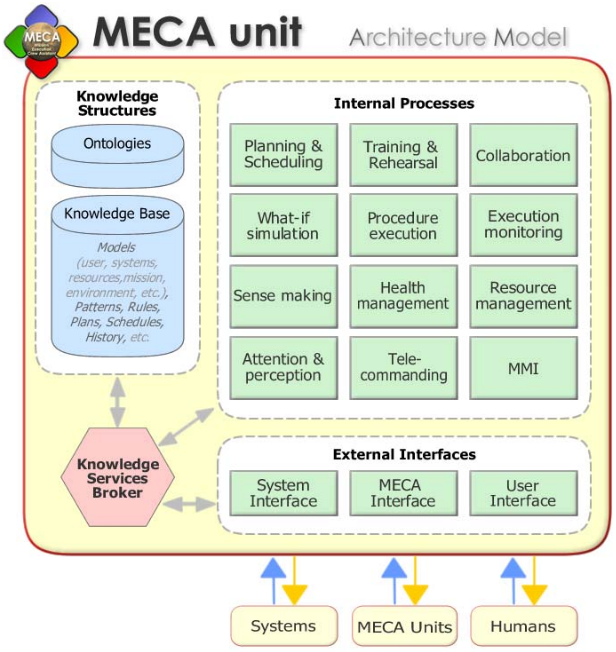


Figure 3: General MECA Architecture.

4.1 User group walkthrough

Four expert engineers in the domain of human exploration and supportive systems participated in the user group walkthrough that took about three hours. The participants were walked through the scenario by a MECA team member and answered questionnaires during and after the walkthrough. Storyboarding was used to illustrate the operations, context and MECA support functions (figure 4).

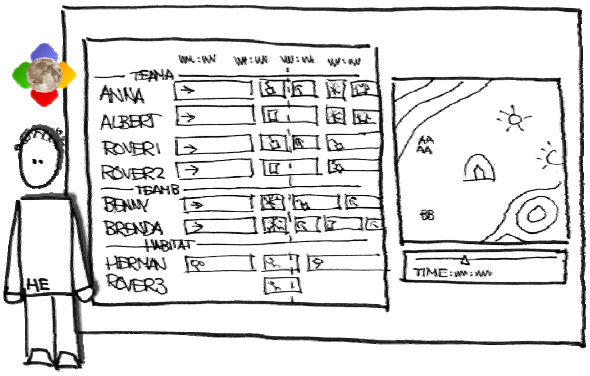


Figure 4: Storyboard picture with Herman in habitat, keeping an overview of all teams on EVA.

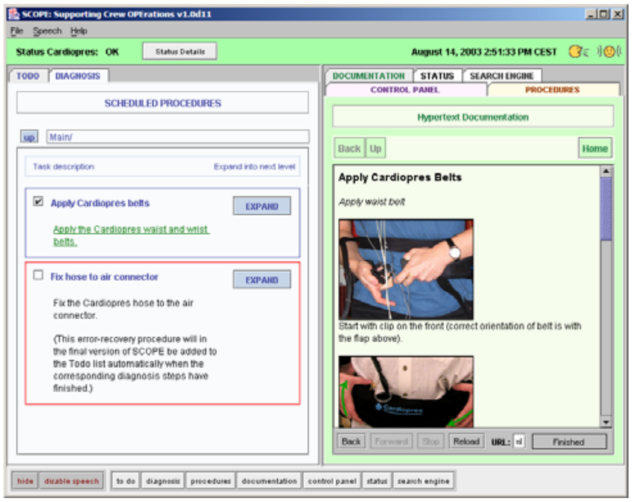
The first objective of the walkthrough was to evaluate the scenario and, if shortcomings are identified, to get recommendations for improvement (adjustment, additions). Important recommendations for improvement were that the scenario should better show that it is an emergency in which all hands are on deck, and that MECA should show alternative solutions. Furthermore, the scenario could be enriched to include remote (MECA-mediated) collaboration, in which the medical specialist (in Team A) is communicating with Herman during the just-in-time training and the task execution itself.

The second objective of the walkthrough was to assess the support functions of MECA in a "scenario-context" and, if shortcomings are identified, to get recommendations to improve MECA's functionality (which should feed into the RB). In particular, the acceptance of a support system that monitors cognitive and emotional state seemed to be low, and it was doubted whether this really could have beneficial effects. Further, it was proposed that MECA could give a confidence level to its suggestions and that it proposes alternatives.

4.2 Human-in-the-loop evaluations

The MECA core functions were included in the prototype incrementally, following a general top-down process in which the support functions are being refined at three abstraction levels. On the first level, based on users' goals and support needs, the system's functions and information provision are specified (i.e. the task level of the user interface is determined). On the second level, the control of the functions and the presentation of the information is specified (i.e. the "look-and-feel" or the communication level of the user interface is established). On the third level, the interface design is implemented in a specific language on a specific platform (i.e. the actual support is established at the implementation level). In the lab-prototype, the focus is on the task level; most functions will be simulated by a human experimenter via the Wizard of Oz method. This prototype allows for human-in-the-loop evaluations, providing support via a state-of-the-art mobile device. It should be noted that the main objective of this prototype is to assess MECA at the task level, and that we did not invest in optimizing the communication level. Most functions of the prototype are simulated ("scripted") via the Wizard of Oz method (i.e., a human plays most of the MECA-functions behind the screens according to a strict protocol), and the environment (Mars surface with objects like rovers and astronauts) is simulated in a virtual environment (projected on the walls).

We conducted a scenario-based evaluation using such a virtual environment (i.e., a Commercial-Of-The-Shelf (COTS) game-engine), [42]. The evaluation consisted of two successive parts: (1) A multi-role test with 11 participants playing the different persons (roles) of the scenario, and (2) a validation test of the results with submarine operators.



☑ Apply Cardiopres belts

Apply the Cardiopres waist and wrist belts.

☐ Fix hose to air connector

Fix the Cardiopres hose to the air connector.

(This error-recovery procedure will in the final version of SCOPE be added to the Todo list automatically when the corresponding diagnosis steps have finished.)

Apply Cardiopres Belts

Apply waist belt

Start with clip on the front (correct orientation of belt is with the flap above).

Figure 5: Example of MECA user interface, providing procedure and diagnosis support.

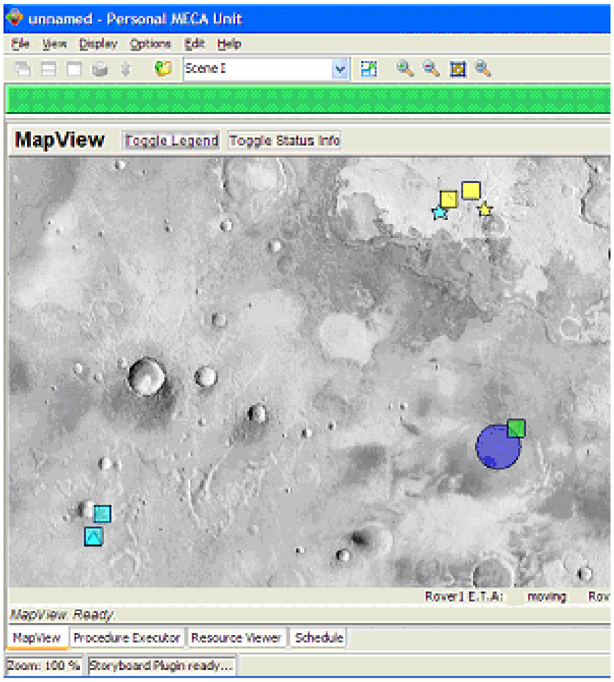


Figure 6: Example of MECA, helping to re-allocate resources (rovers) to cope with critical situations.

In general, the method of game-based evaluation worked well and provided further insight into the user needs for MECA support, in addition to the previous Structured Walkthrough. Distinguishing between task level—where the focus was—and communication level issues proved to be possible. Corresponding to the Structured Walkthrough, participants confirmed the good potential MECA has to support astronauts in the game-based evaluation. Eight out of eleven participants found MECA pleasant to work with. Participants who found MECA unpleasant to work with had comments on the maturity of the prototype, the way of evaluation (no voice communication between participants) and that MECA could be more pleasant if it was resource oriented instead of task oriented. Six out of eleven participants found that MECA offered sufficient support. Participants who disagreed mostly remarked that there should be a single overview page available.

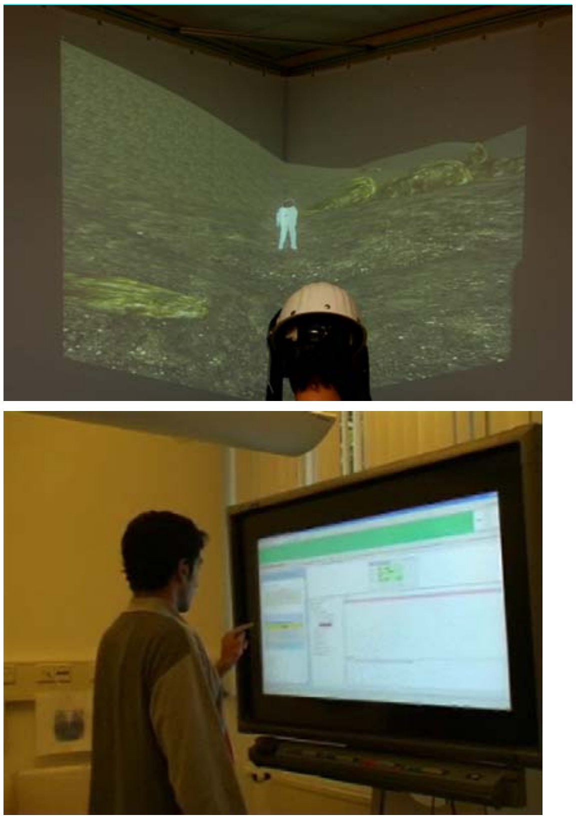


Figure 7: These pictures show two participants playing different roles in the scenario.

Naval operators are highly trained and prepared for nominal and off-nominal situations. Furthermore, they have experience with naval missions that are operated in rather isolation (communication to the "outside" can be blocked), for a relatively long duration (e.g. months), and under extreme conditions (limited space, closed working and living environment). The operations are highly dependent on technology (e.g. energy, oxygen production), and costs of failures are high (i.e., safety is a crucial issue). These specific characteristics of submarine crew and operations show remarkable correspondences with space crew and operations. At this moment, European astronauts were heavily involved in ISS training and could not contribute to our evaluation; therefore, we chose the submarine group to validate the evaluation results. In general, the submarine participants confirmed the outcomes of the first game-based evaluation with 11 participants. They were very positive on the MECA support, and noted that there is a need for such support in the envisioned scenario and environmental conditions. Two observations were of specific relevance. First, the observation that the current version of MECA is too rigid and should better allow for flexible human operations: it should not straightjacket astronauts' work. This should be addressed in MECA's requirements baseline (e.g. on the issue of mixed-initiative) and the next evaluation. Second, the submarine operators emphasized the importance of affective responses: the crew has to adequately cope with such—anticipated—responses, and MECA could provide support on this issue. Dealing with cognitive and affective load should be further worked out in the next prototype.

Whereas the Walkthrough participants expressed reluctance to MECA's function for monitoring astronaut's cognitive load and emotional state, the participants were very positive on this function in the simulation-based evaluation. This may be due to the fact that the last participants did not include "real" astronauts. Based on this result, we will continue to refine this function for the next MECA prototype, and investigate the possible reluctances and ways to overcome them. With respect to the prototype, it is important to allow for more flexibility in the next evaluation. The current Graphical User Interface (GUI) dialogue was rather rigid, providing some hindrance for users to express their support needs during the scenario. A more free speech dialogue, in which speech recognition and synthesis can be simulated in a Wizard of Oz setting, should be included in a next evaluation.

5. CONCLUSIONS AND DISCUSSION

Manned long-duration missions to the Moon and Mars set high operational, human factors and technical demands for a distributed support system, which enhances human-machine teams' capabilities to cope autonomously with unexpected, complex and potentially hazardous situations. Based on a situated Cognitive Engineering (sCE) method, we specified a sound—theoretical and empirical founded—set of requirements for such a system—called Mission Execution Crew Assistant (MECA)—and its rational consisting of scenarios and use cases, user experience claims, and core support functions. We started with an elaborate analysis of (1) past and envisioned missions to the Moon and Mars, (2) current and future technology, that might provide the required operational support, and (3) human factors issues, that appear in such missions and support technology. This resulted in the definition of the distributed personal ePartners concept and a first set of MECA requirements and its rational. This situated Requirements Baseline (RB) was reviewed by persons with relevant but different experiences and expertise. Furthermore, the core MECA functions were incrementally implemented into a lab-prototype (most of the functions were simulated). Subsequently, a human-in-the-loop evaluation was conducted, using a virtual (game) environment.

Table 4: After the first review, we established a Requirements Baseline with 142 requirements. Subsequent evaluations and reviews resulted in a RB with 167 requirements.

Type	Same	Refine	New	Total
Task Level Requirements	16	5	6	**27**
Functional Requirements	27	14	6	**47**
User Interface Requirements	25	15	5	**45**
Technical Requirements	8	2	6	**16**
Operational requirements	1	2	1	**4**
Interface Requirements	27	0	1	**28**
TOTAL:	**104**	**38**	**25**	**167**

MECA comprises distributed personal ePartners that help the team to assess the situation, to determine a suitable course of actions to solve a problem, and to safeguard the astronauts from failures. In addition to standard requirements reviews, we tested and refined the RB via storyboarding and human-in-the-loop evaluations of a simulation-based prototype in a virtual environment. Overall, the sCE method provided a reviewed set of 167 high-level requirements that explicitly refers to the tested scenarios, claims and core support functions on health management, diagnosis, prognosis & prediction, collaboration, resource management, planning, and sense-making (see table 4). Taken together, the evaluation results confirmed the claims on effectiveness, efficiency, satisfaction, learnability, situation awareness, trust and emotion. Based on the results, issues for improvement were identified and prioritized (e.g., acceptance of mental load and emotion sensing, improving team situation awareness, and enriching the speech dialogues).

In general, the situated sCE method proved to work well. First, the method provided a refined and validated scenario, an alternative scenario, and broader set of scenarios and use cases that link to the requirements. Claims were refined and confirmed, and issues for further research prioritized (e.g., acceptance of load and emotion sensing). The participants were positive on the core functions. Future research and developments will have more operational realism (i.e.., the Mars500 programme [43], the International Space Station, and Mars analogue environments on earth): A further development of the MECA knowledge base, which will be tested for long duration, in both off-nominal and nominal situations, in relatively isolation and with more astronaut involvement.

The sCE methodology has been developed and successfully applied for the design of support systems for manned space missions. A recent application of this methodology in the defense domain was also successful: It provided a practical, coherent and extendable requirements baseline for adaptive track-handling support that can be incrementally developed and implemented [44]. It is interesting to note that—for a part—similar human factors theories and models were applied in the work domain and support analysis. By explicating the design rationale in a similar way, it is rather easy to identify which support elements can be applied to the different domains and which elements are really domain specific.

6. ACKNOWLEDGEMENT

We would like to thank all persons who participated in the reviews and evaluations, helping us to improve the MECA RB and prototype. MECA is a development funded by the European Space Agency (Contract Number 19149/05/NL/JA).

7. REFERENCES

[1] Hoc, J.-M. (2001). Towards a cognitive approach to human-machine cooperation in dynamic situations. *International Journal of Human–Computer Studies*, 54(4), 509–540.

[2] Neerincx, M.A., Lindenberg, J., Smets, N., Grant, T., Bos, A., Olmedo Soler, A,. Brauer, U., Wolff, M. (2006). Cognitive Engineering for Long Duration Missions: Human-Machine Collaboration on the Moon and Mars. *SMC-IT 2006: 2nd IEEE International Conference on Space Mission Challenges for Information Technology*, pp. 40-46. Los Alamitos, California: IEEE Conference Publising Services.

[3] Neerincx, M.A. & Lindenberg, J. (2008). Situated cognitive engineering for complex task environments. In: Schraagen, J.M.C., Militello, L., Ormerod, T., & Lipshitz, R. (Eds). *Naturalistic Decision Making and Macrocognition* (pp. 373-390). Aldershot, UK: Ashgate Publishing Limited.

[4] Hollnagel, E., and Woods, D. D. (1983). Cognitive systems engineering: New wine in new bottles. *International Journal of Man–Machine Studies*, 18, 583–600.

[5] Norman, D. A. (1986). Cognitive engineering. In D. A. Norman, and S. W. Draper (Eds) *User-Centered System Design: New perspectives on human–computer interaction*. (pp.31-62). Hillsdale, NJ: Lawrence Erlbaum Associates.

[6] Rasmussen, J. (1986). *Information processing and human–machine interaction: an approach to cognitive engineering*. Amsterdam, Elsevier.

[7] Maanen, P.P. van, Lindenberg, J., and Neerincx, M.A. (2005). Integrating human factors and artificial intelligence in the development of human-machine cooperation. In H.R. Arabnia, and R. Joshua (Eds), *Proceedings of the 2005 International Conference on Artificial Intelligence (ICAI'05)* (pp.10-16). Las Vegas, NV: CSREA Press.

[8] Neerincx, M.A. (2003). Cognitive task load design: model, methods and examples. In: E. Hollnagel (ed.), *Handbook of Cognitive Task Design*. Chapter 13 (pp. 283-305). Mahwah, NJ: Lawrence Erlbaum Associates.

[9] Kanas, N.A., Salnitskiyb, V.P., Ritsher, J.B., Gushin, V.I., Weiss, D.S., Saylor, S.A., Kozerenko, O.P. and Marmar, C.R. (2006). Human interactions in space: ISS vs. Shuttle/Mir. *Acta Astronautica*, 59, 413-419.

[10] Kanas, N.A., Salnitskiyb, V.P., Ritsher, J.B., Gushin, V.I., Weiss, D.S., Saylor, S.A., Kozerenko, O.P. and Marmar, C.R. (2007). Psychosocial interactions during ISS missions. *Acta Astronautica*, 60, 329-335.

[11] Hoffman, S.J. and Kaplan, D.I. (Eds)(1997). Human exploration of Mars: The reference mission of the NASA Mars Exploration Study Team. NASA Special Publication 6107. Houston, TX: NASA Johnson Space Center.

[12] Hoffman, S.J. (ed.)(2001). The Mars Surface Reference Mission: A Description of Human and Robotic Surface Activities. TP-2001-209371, Dec 2001. Houston, TX: NASA Johnson Space Center.

[13] Engel, K., Hirsch, N., Junior, A., Mahler, C., Messina, P., Podhajsky, S. and Welch, C.S. (2004). Lunares: International Lunar Explroation in Preparation for Mars.

Proc. of 55th Int. Astronautical Congress, pp. 1-11. Vancouver, Canada

[14] Kminek, G., (2004). Human Mars Mission Project: Human. surface operations on Mars, ESA/Aurora/GK/EE/004.04,. issue 1, revision 1, June 2004. Noordwijk, The Netherlands: ESA Publications Division.

[15] HUMEX (2003). HUMEX: A study on the survivability and adaptation of humans to long-duration exploration missions. ESA Special Publication 1264. Noordwijk, The Netherlands: ESA Publications Division.

[16] Baker, C., Naikar, N. and Neerincx, M.A. (2008). Engineering planetary exploration systems: integrating novel technologies and the human element using work domain analysis. *Proceedings of the 59th International Astronautical Congress (IAC2008)*, Glasgow.

[17] Vicente, KJ (1999). *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer Based Work*, Lawrence Erlbaum Associates, London.

[18] Naikar, N & Sanderson, P.M. (2001). Evaluating Design Proposals for Complex Systems with Work Domain Analysis, *Human Factors*, 43(4), 529-542.

[19] Endsley, M.R.. (2000). Theoretical Underpinnings of Situation Awareness. In M.R. Endsley & D.J. Garland (eds). *Situation Awareness Analysis and Measurement.* LEA, Mahwah, NJ, USA.

[20] Weick, K. (1995). *Sense-making in Organisations*. Sage, Thousand Oaks, CA, USA, 1995.

[21] Klein, G.. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.

[22] Scerbo, M.W. (2001). Stress, workload and boredom in vigilance: a problem and an answer. In Hancock, P.A. & Desmond, P.A. (ed.), *Stress, Workload and Fatigue*. Mahwah, New Jersey, Lawrence Erlbaum Associates.

[23] Adams, B.D., Bruyn, L.E., Houde, S. and Angelopoulos, P. (2003). *Trust in Automated Systems: Literature Review*. DRDC Report No. CR-2003-096. Toronto, Canada: Defence Research and Development.

[24] Neerincx, M.A. (2007). Modelling Cognitive and Affective Load for the Design of Human-Machine Collaboration. In: D. Harris (Ed.). *Engineering Psychology and Cognitive Ergonomics, HCII 2007*, LNAI 4562, pp. 568–574. Berlin Heidelberg: Springer-Verlag. ISBN 978-3-540-73330-0

[25] Mohammed, S. & Dumville, B.C.. (2001). Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behaviour*, 22(2), pp. 89-106.

[26] Helmreich, R.L., Merritt, A.C., & Wilhelm, J.A.. (1999). The evolution of Crew Resource Management training in commercial aviation. *International Journal of Aviation Psychology*, 9(1), pp. 19-32.

[27] Aarts, E., Harwig, R. & Schuurman, M. (2001). Ambient Intelligence. In P.Denning (Ed.), *The Invisible Future* (pp. 235-250). New York: McGraw Hill.

[28] Lindenberg, J. Pasman, W., Kranenborg, K., Stegeman, J. & Neerincx, M.A. (2007). Improving service matching and selection in ubiquitous computing environments: a user study. *Personal and Ubiquitous Computing*, 11, 59-68.

[29] Satyanarayanan M (2001) Pervasive computing: vision and challenges. *IEEE Pers Commun* 8(4):10–17.

[30] Dowding, J., Alena, R., Clancey, W. J., Graham, J., and Sierhuis, M. (2006). Are you talking to Me? Dialogue Systems Supporting Mixed Teams of Humans and Robots. A*AAI Fall Symposium 2006: Aurally Informed Performance: Integrating Machine Listening and Auditory Presentation in Robotic Systems*, October, Washington, DC.

[31] Fogg, B.J. (2003). *Persuasive technology: Using computers to change what we think and do*. Amsterdam etc: Morgan Kaufmann Publishers.

[32] Fong, T., Nourbakhsh, I. & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42, 143-166.

[33] Gajos K, Fox H, Shrobe H (2002) End user empowerment in human centered pervasive computing. In: *Proceedings Pervasive 2002*, Zurich, Switzerland, pp 134–140

[34] Goetz, J., Kiesler, S. & Powers, A. (2003). Matching robot appearance and behavior to tasks to improve human-robot cooperation. *IEEE Ro-Man* 2003, 55-60.

[35] Looije, R., Neerincx, M.A., and de Lange, V., (2008). Children’s responses and opinion on three bots that motivate, educate and play. *Journal of Physical Agents*, vol 2(2), 13-20.

[36] Clancey, W.J. (2004). Roles for Agent Assistants in Field Science: Understanding Personal Projects and Collaboration. I*EEE Transactions on Systems, Man and Cybernetics—Part C: Applications and Reviews*, Vol. 34, No. 2, May 2004.

[37] Clancey, W.J. Sierhuis, M., Alena, R., Berrios, D., Dowding, J., Graham, J.S., Tyree, K.S., Hirsh, R.L., Garry, W.B., Semple, A., Buckingham Shum, S.J., Shadbolt, N., and Rupert, S.M. (2005) Automating CapCom Using Mobile Agents and Robotic Assistants. In Proceedings American Institute of Aeronautics and Astronautics 1st Space Exploration Conference, 31 Jan - 1 Feb, 2005. [Available from: AIAA Meeting Papers on Disc [CD-ROM]: Reston, VA, and as AKT-IRC ePrint 375: http://eprints.aktors.org/375], Orlando, FL

[38] Hirsh, R., Graham, J., Tyree, K., Sierhuis, M., and Clancey, W. J. (2006) Intelligence for human-robotic planetary surface robots. In A. M. Howard and E. W. Tunstel (Eds.) *Intelligence for Space Robotics*, Albuquerque: TSI Press, pp. 261-279.

[39] Dong, S. et al. (2008). Self-assembling wireless autonomously reconfigurable module design concept. Acta Astronautica, 62(2-3), 246-256

[40] Rosson, M. B. and Carroll, J. M. (2001). *Usability engineering: Scenario-based development of human-*

computer interaction. San Francisco, CA: Morgan Kaufman.

[41] McCrickard, D.S. and Chewar, C.M. (2006). Designing Attention-Centric Notification Systems: Five HCI Challenges. In: J. Chris Forsythe, Michael L. Bernard, and Timothy E. Goldsmith (Eds.). *Cognitive Systems: Human Cognitive Models in Systems Design*, pp. 67-89. Lawrence Earlbaum

[42] Smets, N.J.J.M., Abbing, M.S., Neerincx, M.A., Lindenberg, J, and van Oostendorp, H. (2008). Game-based evaluation of personalized support for astronauts in long duration missions. *Proceedings of the 59th International Astronautical Congress (IAC2008)*, Glasgow.

[43] Rauterberg, M., Neerincx, M.A., Tuyls, K., and Loon, J. van (2008). Entertainment computing in the orbit. In: Paolo Ciancarini, Ryohei Nakatsu, Matthias Rauterberg, Marco Roccetti (Eds.). *IFIP International Federation for Information Processing, Volume 279; New Frontiers for Entertainment Computing*; pp. 59–70. Boston: Springer.

[44] Neerincx, M.A., Te Brake, G.M., Van de Ven, J.G.M., Arciszewski, H.F.R., De Greef, T.E., and Lindenberg, J. (2008). Situated cognitive engineering: Developing adaptive track handling support for naval command and control centers. In: B. Patrick, C. Gilles and L. Philippe (Eds.). *HCP-2008 - Third Internation Conference on Human-Centered Processes*, pp. 3-20. Bretagne, France: TELECOM Bretagne ISBN: 978-2-908849-22-6.