Simulating "Mars on Earth"
A Report from FMARS Phase 2

William J. Clancey

Institute for Human and Machine Cognition
University of West Florida, Pensacola
bclancey@arc.nasa.gov

Abstract
By now, everyone who's heard of the Haughton-Mars Project knows that we travel to Devon Island to learn how people will live and work on Mars. But how do we learn about Mars operations from what happens in the Arctic? We must document our experience—the traverses, life in the hab, instrument deployment, communications, and so on. Then we must analyze and formally model what happens. In short, while most scientists are studying the crater, other scientists must be studying the expedition itself. That's what I have done in the past four field seasons. I study field science, both as it naturally occurs at Haughton (unconstrained by a "Mars sim") and as a constrained experiment using the Flashline Mars Arctic Research Station.

During the second week of July 2001, I lived and worked in the hab as part of the Phase 2 crew of six. Besides participating in all activities, I took many photographs and time lapse video. The result of my work will be a computer simulation of how we lived and worked in the hab. It won't be a model of particular people or even my own phase per se, but a pastiche that demonstrates (a proof of concept) that we have appropriate tools for simulating the layout of the hab and daily routines followed by the group and individual scientists. Activities—how people spend their time—are the focus of my observations for building such a simulation model.

The FMARS Simulation
The FMARS simulation is constructed using a tool called Brahms (Clancey et al. 1998; Sierhuis 2001), which we are developing at NASA/Ames Research Center. The components of a Brahms model are fairly easy to understand:

- People (Agents & Groups, e.g., biologists, the Phase 2 crew, the Capcom role)
- Geography (the building and its layout)
- Objects (e.g., cameras, tables, suits, documents)
- Activities (e.g., reading email, EVA prep, watching movie, downloading & sharing digital photos, debriefing, sleeping, waiting for help to remove suit)

Why do we want to build such a model? Wouldn't it be sufficient to write an ethnographic report? Most importantly, model building is a tool for developing better social-psychological theories of human behavior. Through building the FMARS model, we have come to better understand the nature of joint activities (collaboration, e.g., filling the water tank) versus group

1 On leave at NASA/Ames Research Center, Mountain View, CA.
2 Boris Brodsky (QSS) is the primary model builder; Maarten Sierhuis (RIACS) is the Brahms Project Manager; Ron vanHoof (QSS) is the Brahms System Manager; Mike Scott (QSS) is a Senior Programmer.
activities (working together, but independently, e.g., using computers), the nature of *dynamic interactions* (e.g., following someone) versus planned actions, and the different *motives* (e.g., having fun, physiological needs) behind purposeful activity (Clancey, in preparation).

We also simulate life in the hab rather than just describing it because other aspects of the hab, especially the life support systems and controlling software, will be extensively formalized and simulated in computer programs as part of a design and test process. Without a complementary model of human behaviors, these system simulations will make assumptions about loads placed by human activities (e.g., the power required in the hab at different times). System simulations also tend to use simple models of how people use an interface to command control adjustments or respond to control software requests. Using Brahms, we can develop an integrated simulation of systems and human behavior.

Finally, during a Mars mission itself a simulation could be useful for testing and instruction of revised procedures. For example, we could revise a simulation to illustrate a new procedure, perhaps using new systems software, and then upload the procedure, software, and integrated simulation to a Mars crew for them to investigate and interact with (i.e., real people interacting with simulated systems and agents). This could provide confidence that the new design will work, as well as providing a valuable tool for eliciting the crew’s comments. With the inability to converse directly with the crew, the support on Earth, we could even send simulated agents for the crew to interact with, agents who convey the models and explanations of specialists back on Earth.

To summarize, the applications of a hab simulation include:

- Habitat design
- Automation design and testing
- Formalizing analog experimental protocols
- Crew scheduling
- Communication-coordination planning
- Training (especially by interacting within a simulation)
- Education – public outreach
- Research on work and behavior modeling

Our immediate interest is to develop Brahms well enough so the various applications can be explored in research projects. For example, through NASA funding we have integrated the FMARS simulation with an existing simulation of an air recycling system and an artificial intelligence monitoring and control system (Malin et al. 2000). The FMARS simulation will place loads on the recycling system, providing a contextual model of hab operations for testing the AI software. Furthermore, the (simulated) crew will interact with the AI software, for example, getting information about resource capacity (e.g., oxygen reserves) needed for planning daily work. Applying the methods of instructional systems (e.g., Clancey & Soloway 1990), we could develop programs that use a Brahms model to understand what the crew is doing, so the programs could provide appropriate support.

**Methods**

How do we build the FMARS model? There are two primary methods (Clancey 2001): *Participant observation* (learning by being a member of the crew) and *photographic documentation* (including time lapse). During my week in the hab, I took regular notes about who did what, where, when, and why. Each day I added to this, refining with details, and finally developing hypotheses about why activities unfold in the manner I observed. In short, we need a theory of “what happens next.” What determines the next behavior of individuals and the group?
To organize my observations, I created a table in a document, with columns for the name of the activity, the location where it occurred, the time, who participated, and comments. For regular activities, such as EVAs and meetings, I used the table to record when the activity began and ended. By the fourth or fifth day I was able to sort the table more or less chronologically for a typical day and segment it into broader categories (e.g., breakfast, briefing/planning, EVA). Towards the end of the week, I began to refine some activities into subcategories (e.g., reasons for working at a laptop). Finally, after I left the hab, I realized the significance of activities and modes of behaving that I had not thought to write down earlier (e.g., listening to music while working at the computer).

My other notes were organized in an outline of incidents and issues, as they emerged during my stay:

- Steps in an EVA
- Troubleshooting incidents (e.g., electric generator, space suits)
- Why staterooms are not used during the day
- Traverse planning and navigation (e.g., avoiding mud)
- Problematic spaces (e.g., the workstation area is cramped)
- Non-issues (e.g., staterooms are quiet)
- Learning from analog experiments; how to improve them

At various times I wrote down where everyone was in the hab and what they were doing. This provides a snapshot of life in the hab ("snaplists"). In retrospect, I should have done this on a regular basis (e.g., once an hour), for it would be a good way of verifying the simulation model. I had also intended to follow someone every day, to note their behaviors in some detail, but as a participant in the hab, where group activities dominated (mostly organized around EVAs), this proved impractical. Finally, after I began to understand why activities occurred when they did, I realized I needed statistical information about events (e.g., how often and when we received radio calls from base camp). Some of this information should have been logged by the crew (e.g., generator failures). Other information, such as external communications, could have been logged by mission support.

It requires more than a week to realize all that one might study, especially if psychosocial factors are included. I believe that several weeks would be necessary to realize what categories are relevant; in general, multiple stays with different crew combinations are desirable for making generalizations and understanding crew-specific practices.

Results
What are the results of my observations? I now have a table with about fifty activities, grouped according to broad "times of the day." Here is an initial description of these broad periods during a day in the life of FMARS 2001 Phase 2:

- 0700-0900 Breakfast
- 0900-1030 Briefing/Planning
- 1030-1500 EVA
- 1500-1530 Eat and Clean up
- 1530-1700 Briefing and planning
- 1700-2000 Computers (email, photo download, software testing); data analysis in lab
- 2000-2100 dinner and cleanup
- 2100-2400 movies, refreshments (especially chocolate)
- after midnight: sleep, reading and writing
This outline is a broad abstraction, an average of seven days, not a schedule we followed. Nevertheless, the patterns can be striking. For example, on three sequential days the EVA crew stepped into the airlock at 1105, 1106, and 1108. No procedure required that we do this, it was just an emergent product of our intentions, the constraints of getting into suits and fixing radios, and our other habits (such as when we awoke, how long it takes to eat, and time to arrange personal gear). Absolute times will vary each day, but relative times, such as when a debriefing occurs after an EVA, are more regular (in this case, about 30 minutes). This chaining of group activities is a key part of the order of the day (which might be explained as part of individual, psychological processes).

What I have said so far should make clear why it's not reasonable to expect a "human factors" report from the hab every day, providing research results. Unlike the biologists and geologists, I do not collect isolated samples in plastic bags. My daily observations are mostly too mundane to mention (as the pattern itself hardly seems surprising). Also, it takes four to five days until apparent habits are established, and then a few more days before details can be filled in (e.g., what are people doing for so many hours at their computers?).

**Time Lapse Video Example of EVA Planning**

An example analysis of a time lapse video reveals how I do my work and what can be learned. Based on an experiment in the initial year 2000 occupation of the hab, I placed my Hi-8 video camera in the far corner, in front of the right-most stateroom near the SE portal. I captured quarter-frame images (320 x 240 pixels) direct to disk every 3 seconds using a PC Card and video-editing program. Experience in analyzing such time lapses since HMP-1999 showed this frequency to be useful and sufficient. I captured two entire days in this way. In retrospect, I might have left the time lapse running every day. Full analysis is tedious, but the time lapse is also useful for capturing broadly the behavior of the group during critical periods, as the following example illustrates.

Before the EVA of July 15, the group discussed where to go. I was not part of the EVA crew, so I sat to the side writing, but also observing. When the group all gathered around the commander's laptop, I began paying more attention and took photographs of the ensuing action. Similarly, the Discovery photographer picked up his video camera and began filming the action. The incident was immediately interesting because it illustrated a group planning activity, using multiple representations, coordinated with views through the portal, with people pointing and calling attention to features throughout.

Fortunately, we have a time lapse recording of this activity, so we can see all movements, who is initiating changes, and when the changes occur (within three seconds). The group moves like a flock of birds during a 12 minute period, gathering at the laptop Landsat image (Figure 1), NW Portal (Figure 2), a projected map of crater, and an air photo on table. Everyone participates. The commander tends to move in broad sweeps from one end of the floor to the other, to be contrasted with pivoting around the central area of the floor or tagging along. One person appears to be interacting with the commander most closely in these movements, suggesting a joint decision-making process. Another crew member lags behind with his coffee, but always joins the group to share their view.
Strikingly, the activity is clearly over when the commander, the crew member who was speaking with him throughout, and two other crewmembers stand to form a square and laugh. It looks like closure (Figure 3). The group then obviously disperses to prepare for the EVA.

This example illustrates the value of having a time lapse recording at all times for the sake of capturing such group activities. The photographs also illustrate that the conventional manner of documenting such activity (notice the Discovery Channel cameraman to the right of the group in Figure 1) fails to capture the overall pattern of how people are gathering and moving as a group. The time lapse shows very well how one or two people reorganize the activity by calling attention to different representations (the photos and maps) and the views out the window. Thus, the use of representations (including of course people's utterances) is strongly contributing to individual attention, such that we can talk about a group activity. Finally, the time lapse provides a means of quantifying the duration and phases of the decision-making activity. The end of the activity is particularly well marked (Figure 3).
How might this understanding of group planning be useful? Consider the value to mission support on Earth of such (time-delayed) video. If the group weren’t moving together, we might wonder whether there was a disagreement. Or perhaps the group had broken into subgroups to plan more quickly. Thus knowing what group planning looks like under different conditions would be a useful clue to mission support about the crew’s attitudes. A theory of dialog as a spatial phenomena, not limited to a computer screen, would also be useful for designing robots that track human speech, gaze, and gestures to understand our intentions and communicate with us.³

**Layout and use of space**

An important part of the Brahms simulation of FMARS is a virtual reality depiction of the facility. The data gathered includes extensive photographs of all objects and areas, close-up photographs for color and texture rendering, and a scale drawing of the hab (Figure 4). This drawing shows the layout at a particular time, with the precise arrangement of laptops and chairs.

Figure 5 illustrates how Brahms simulations will appear on the computer screen using a browser. Although we call this the Brahms-VR system, the viewer is not immersed in the world, as in a virtual reality system. We refer to it as a 2 1/2d depiction. The viewer could appear on-screen as an “avatar,” a crew member in the simulation. In the prototype implementation, we use the Brahms-VR system to view a simulation after it completes. The target design involves dynamic interaction, so one may view the simulation while it is running, and the simulation itself will draw upon physics modeling in the rendered world (e.g., to determine the path taken by an agent to avoid collision). There are many fascinating problems to be solved here, including simulating agent postures, loudness of sounds (e.g., can someone on the upper deck hear a conversation below?), sightlines, and joint actions (e.g., two people carrying an object).

³ Randy Stiles, personal communication.
Figure 4. Drawing of FMARS upper deck at 16:15 July 13, 2001. One chair is on the lower deck; accuracy of laptop locations and chairs etc. within 3".

Figure 5. Snapshot of virtual reality depiction of upper deck.
Of special interest is how use of hab facilities varies over time and with different activities (cf. Kantowitz & Sorkin 1983). Photographs and other observations suggest that the layout of tables is perhaps the most important aspect of space design in the upper deck, aside from the staterooms and storage areas. The workstation area is the most obvious area where design requires improvement. The built-in table is not deep enough (about 24 inches) and is too cramped for six laptops plus a large server display (which hogs the most attractive area below the portal and blocks the view). The mess table provided extra space (Figure 6).

Figure 6. Using the table for a joint task (filling out a questionnaire)

Figure 6 shows a group activity, in which people are gathered together. In contrast, Figure 7 shows the activity of individual work; people are apparently maximizing their separation.

Figure 7. Individual work, maximizing spacing from each other.
Compare Figure 8, showing a different layout used by another FMARS crew, later in the field season.

![Figure 8](image)

Figure 8. Alternative configuration in a different phase; separating the work area from the galley.

Now it isn't possible to sit around one table as a group (this crew went to base camp for dinner). More work area is provided, relieving the elbow-to-elbow crowding of the curved workstation area along the wall. Consistent with this configuration, the tables on the upper deck were used as a general workspace area during Phase 2. Besides eating here, people used the tables to: Play games, look at maps, clean cameras, prepare instruments for deployment, assemble and test communications equipment, and analyze data (e.g., the magnification apparatus on the right table in Figure 8). Notice also how items are stored on the tables during the day, including cups, notebooks, samples, etc. Would samples be brought into the upper deck on Mars? Why aren't equipment-related activities occurring downstairs? Issues of space, lighting, and room temperature need to be considered (the lower deck of FMARS was typically 5 degrees C cooler). Perhaps also the group enjoys being together.

**Activity Drivers: What determines what people do next?**

The most detailed aspect of the Brahms simulation is a description of each activity as a set of conditional steps or alternative methods. That is, the conditions—when an activity is performed—must be specified. Given the table of activities (outlined above), we see that *group activities are the main driver of behaviors* in the hab, fitting the chronology of the day: Breakfast, Briefing, EVA, Debriefing, Dinner, and Movie. That is, during this phase in the hab, individual behavior is constrained most strongly by coordinated group interactions. Furthermore, the daily EVA is the central, pivotal activity of the day, with meetings, preparations, and even meals occurring around it. This implies that the backbone of the simulation will be behaviors individuals inherit (in the Brahms representation) from the “Hab Crew” group. Each behavior in Brahms is represented as a *workframe*, which is a situation-action rule. In general, the situation (conditional part) of the key workframes for Hab Crew activities will specify either the time of day (e.g., morning briefing) and previously completed activities (e.g., the post-EVA briefing).

*Interruptions* are secondary driver of behavior, including: Radio calls (from base camp) or satellite phone calls (usually pertaining to our communications systems), systems emergencies...
(toilet, comms), **hab maintenance** (refilling the water reservoir, refilling the generators), and **media interviews** (conducted in the lower deck). Frequency information for the radio and phone calls might be determined from the time lapse. I did not have the time (or presence of mind) to systematically gather information about the frequency and timing of when these activities occurred.

*Individual activities,* behaviors that are individually motivated and performed alone, fill the remainder of the day:

- science data processing (e.g., analyzing dosimeter data)
- report filing (both individual reports and the daily crew report prepared by the commander)
- email
- digital picture processing (backup, sorting, sharing),
- chores (cooking, emptying garbage, etc.)
- personal hygiene
- taking photos inside the hab (personal documentation)
- recreational reading

In summary, the **conditions** on activities are the group's practice, interruptions (reactive behavior), and individual practice. Individual activities may be **periodic** (e.g., checking email), **based on time** (a crew or personal habit) or based on **remembering** something you planned to do in the hab. Group practice is mostly **chronological**, but is also **scheduled** as required (e.g., cleaning the suits), **chained** (briefing follows an EVA), and **reactive** (e.g., handling emergencies such as an electrical short in a backpack). Further analysis of this classification has led to better articulating the nature of activity model, as compared to activity theory and task analysis (Clancey in preparation).

**Lessons Learned about Analog Research**

Of paramount importance, given the effort to build the hab, is determining what can be learned from an analog activity such as FMARS and how the activity should be managed and controlled as a scientific investigation. As incidents occur, such as problems with the generator and getting stuck in the mud during a traverse, one naturally realizes that the similarities and differences to Mars must be sorted out. Here is one breakdown:

- What is relevant to the Arctic only? E.g., refueling generators with gas
- What might occur for different reasons on Mars? E.g., getting stuck in sand instead of mud
- What's important that we might do better? E.g., interactions with mission support
- What's difficult on Mars but easy in Arctic? E.g., removing helmet to clean visor

In general, the operation of FMARS suggests a tension between **authentic science** (with mission support as required) and **simulated operations** (on the surface and at mission support). **Taken to the extreme,** the first point of view is that FMARS is a research station to be used as a place for living and working in the Arctic, carrying out studies of science and robotics that are appropriate for Mars mission planning. The other point of view is that FMARS is first of all a simulated hab and crew occupations are simulated missions to Mars. In a simulation, there must be some clearly defined procedures and a notion of "breaking the sim" (e.g., to clean a visor caked with mud during an EVA). This notion is not relevant to the first idea, in which FMARS is just a habitat.
Is a compromise possible? Here is one suggestion: Analog habitats should be managed primarily for authentic science, as a real mission operation (with planning & training), and secondarily as a research vehicle that simulates Mars (e.g., integration of hab, surface, and support operations). If the activities are not authentic, then the simulation has no grounding. Alternatively, if the activities are authentic, then if the simulated constraints are limited but still valid, we will still learn about scientific exploration relevant to Mars.

Authenticity in mission operations entails documenting communication and coordination protocols in advance, and training everyone in simulations prior to the field season. That is, living and working in the hab is viewed as being a mission, not a sim. (It might be possible to do the sim on-site at the Mars Desert Research Station, prior to the formal occupation.)

**Example of analog analysis**

Abstracting lessons from incidents is not easy. Often the literal events are irrelevant, but a broader moral lies in the taken-for-granted context in which the event occurs. An incident during crew planning for an EVA illustrates my point.

The literal events are obvious to the observer: The hab crew is standing around the wardroom table, discussing how to set a GPS device for the planned von Braun Planitia EVA. Which GPS “system” should they use? The discussion concerned the nature of a GPS measuring system, and revealed that there were two alternatives available on their GPS devices (WGS84 [degree latitude and longitude] and UTM [metric distance]). One crewmember claimed that UTM was becoming standard; the people going on the EVA were more familiar with measurements in degree/minutes. They didn’t know how to use the GPS device to use or read UTM measurements.

The lesson to be drawn is not that astronauts going to Mars should be trained on how to use their equipment or that standards should be adopted before the mission begins. Everyone already knows this. Surely a funded, actual mission would have prepared the crew better. Rather, the less salient and important issue is that route planning was occurring just before the EVA, not the evening before as I observed in HMP-1998 or as was deliberately scheduled during HMP-1999 to coordinate with a mission support team in Houston. If mission support personnel were involved in choosing routes, then the crew wouldn’t be allowed to wait until just before departure to plan the EVA, this would have to be done the previous evening. So why did this group develop a different practice? My observation is that we were too tired from the day’s activities and too busy reporting what we had done to think about the next day. Thus, we may have a real issue here, which perhaps we thought we had understood in 1999. (Also, the 1999 after-dinner communication with Houston focused more on reporting than planning.)

The example illustrates how a simple incident had to be interpreted in the context of previous field seasons, with background knowledge about past NASA practices and expectations. From this we see that FMARS provides a research opportunity for communications research, which can be exploited by enlisting more collaboration and establishing more formal protocols for hab activities (e.g., working more rigorously according to a schedule that is coordinated with Earth operations). Do we have the funds and committed external organizations (e.g., NASA, the Mars Society) to provide a realistic mission support role? Or should communications research using FMARS be focused more on interactions between the EVA crew and the hab?

The GPS example also illustrates that superficial reports about FMARS operations are unlikely to give the viewer an understanding of what we are actually learning or could learn from future analog experiments. Such analysis is especially the province of participant observers and modelers who study the scientists and operations in the analog setting, applying the methods and perspectives like those I present in this report.
FMARS Research Opportunities

What scientific research can be done at FMARS? Researchers may want to consider this list when making proposals for participating in hab activities:

- **Communication Protocols**
  - Enforce and document communications between EVA crew and mission support. Investigate especially exchange of contextual information and instructions for using equipment.
  - Contrast: Commander serves as communications hub with world’s experts/advisors vs. specialists in the crew individually communicate with their own peers and private contacts
  - Improve communications between the EVA crew and hab support using radio and video to provide a running commentary

- **Computer Infrastructure**
  - Integrate computers used for data gathering and analysis with the hab’s communications and computing system
  - Facilitate and study data sharing among the crew (e.g., exchanging files using compact flash cards vs. a shared network)
  - Develop computer records that crews leave behind for subsequent crews: Photographs, articles, history of activities, maps (should mission support provide or supplement this repository?)

- **Laboratory and Data Analysis**
  - Determine the adequacy of laboratory equipment in the hab for data collection and analysis (e.g., rock slicing, microscope with camera)
  - Work collaboratively with experts on “Earth” for data collection planning, analysis, and interpretation.
  - Determine laboratory space required and how to prevent intrusion from other activities

- **Living and Working Priorities in the Hab**
  As a rough cut, one can order priorities for the crew’s attention, based on basic needs and their interactions:
  - Electricity, Toilet (electric), Water (pumped from outside hab), Food
  - Private and quiet areas, especially dry and well-ventilated staterooms
  - Unscheduled time for sleep and individual activities
  - Work areas with adjacent space for personal items (e.g., notes, drinks)
  - Personal storage areas (e.g., for cameras to be ready at hand)
  - Dry and warm EVA clothing (suits)
  - Cleanliness (showers, hot water)
  - Entertainment (e.g., DVD movies)

Possibly the only item in the list of living and working priorities that might differ from previous studies of expeditions (e.g., Stuster 1996) is the relative priority for unscheduled time. Because of good weather opportunities, interest in trying the suits, presence of the press, and shortened phase duration it was desirable to have a significant EVA each day. Preparation and reporting filled most of the remaining time, so the crew was far from being bored or feeling confined. A pre- and post-occupation survey related to this list would be useful.
Using Brahms, we could formalize different schedules, layouts, and resource decisions (e.g., use of water). Design of space habitats and missions will likely involve many tradeoffs and compromises, which a comprehensive simulation should enable us to describe and evaluate.

Acknowledgments
I am grateful to the members of the FMARS Phase 2 crew for participating in this research: Steve Brabam, Charles Cockell, Vladimir Pletser, Katy Quinn, and Robert Zubrin. Conversations with Pascal Lee (Principal Investigator of the Haughton-Mars Project), Maarten Sierhuis, Mike Shafto, and members of the Brahms team have been especially valuable. For more information about Brahms see http://www.agentisolutions.com and papers at http://wjClancey.home.att.net. Funding for this research has been provided by the NASA’s Intelligent Systems Program, Space Human Factors Engineering, and University of West Florida. The virtual world representation of the hab has been developed by Bruce Darner and his associates at Digital Space, Inc. under NASA-STTR funding. See www.digitalspace.com for information about the Adobe Atmosphere implementation.

References