A CLOSED MARS ANALOG SIMULATION: THE APPROACH OF CREW 5 AT THE MARS DESERT RESEARCH STATION

William J. Clancey

NASA-Ames Research Center
Computational Sciences Division MS269-3
Moffett Field, CA 94035
and
Institute for Human and Machine Cognition
University of West Florida, Pensacola
bclancey@arc.nasa.gov

ABSTRACT

For twelve days in April 2002 we performed a *closed simulation* in the Mars Desert Research Station, isolated from other people, as on Mars, while performing systematic surface exploration and life support chores. Email provided our only means of contact; no phone or radio conversations were possible. All mission-related messages were mediated by a remote mission support team. This protocol enabled a systematic and controlled study of crew activities, scheduling, and use of space. The analysis presented here focuses on two questions: Where did the time go—why did people feel rushed and unable to complete their work? How can we measure and model productivity, to compare habitat designs, schedules, roles, and tools? Analysis suggests that a simple scheduling change—having lunch and dinner earlier, plus eliminating afternoon meetings—increased the available productive time by 41%.

HIGHLIGHTS: PROTOCOL AND ACTIVITIES

The fifth crew[1] occupied the Mars Desert Research Station (MDRS) near Hanksville, Utah for two weeks in April 2002. This facility is very similar to the Flashline Mars Arctic Research Station (Clancey, 2000b, 2001b) designed as an analog of a Mars surface habitat. MDRS includes six private staterooms, galley, workstations, and meeting/eating area on the upper deck, with a laboratory, toilet, shower, and EVA preparation rooms on the lower deck.

After a few hours of overlap with the previous crew for orientation, we devoted the first day to setting up equipment and working with Mars Society volunteers to improve the station. The second day, Monday, April 8, began a twelve day period of the formal simulation. During this period, we were alone, except for two short visits by a contractor resupplying fuel and water. All other conversations with humanity were restricted to email, with all mission-oriented (non-personal) communications restricted further to a single point of contact (a member of mission support staffed by the Northern California Mars Society) serving as capcom ("capsule communicator," a NASA term stemming from the Mercury program of the 1960s). All reports, requests, and assistance was first directed through capcom. By protocol, after a secondary

[1] With the author serving as the commander, the MDRS5 crew included Andrea Fori, Jan Osburg, Vladimir Pletser, David Real, and Nancy Wood.

contact was established (e.g., someone to advise our work on the greenhouse), further conversations on the same topic were not mediated by capcom, but were always copied to him.

We organized two special activities:

- A simulated multiple-failure EVA situation (Appendix II).
- A full-day open house for the international press. (German: ARD TV, RTL TV, and Der Spiegel; American: TechTV, Phoenix Fox-10 TV, Swiss: FACTS; Norwegian: Dagbladet Daily; and British: Sunday Telegraph of London).

Consequently the two-week simulation period had three special days before the Sunday of departure: a day of rest (when we conducted the simulated EVA failures), the open house on Saturday before departure, and the clean-up day before the open house. In sum, there were ten actual simulation days, Monday-Friday of the first week and Saturday, plus Monday-Thursday of the second week. This distinction turns out to be very useful for analyzing the group's behavior.

The crew investigated the following issues during their stay:

- If there is life on Mars, how do you take a soil or rock sample that includes it?
- "Expedition memory": Can a geologist understand the work performed by previous rotations to develop a geology primer of the region?
- What is the effect of chores (life support maintenance) on science productivity?
- What are the psychological effects of growing plants in the hab?
- How do plans develop and change during the mission? How do individual and group activities interact?
- How can Earth's mission support understand and assist Mars surface exploration? Can possible EVA targets and routes be suggested using reports from previous crews?
- How is public and private space used? How can the habitat's layout be improved?

METHODS AND DATA ANALYSIS

To address questions about productivity, planning, and layout, the following data were collected:

- Time lapse video of upper deck throughout the rotation (every 3 second, 320x240 pixels) from 730 am until midnight (or on several occasions until everyone is asleep)
- Video recording with sound of all planning meetings
- Log of crew location and activities every 15 minutes on two consecutive days ("snaplist")
- Personal crew logs of awake and sleep times, plus time devoted to being the cook for the day ("director of galley operations" or DGO).
- Written plan of daily proposed and deferred tasks in a table by date and person
- Written crew ("post-occupation") surveys
- 97 reports posted on the web, with completion dates, including commander's daily log and health and safety officer's (HSO) daily reports
- Approximately 1000 time-stamped digital photographs

Effect of scheduling on productivity

To address the question of "where does the time go," we need to consider what time is available, subtracting sleep time, group activities (especially meals and EVA operations), and unscheduled interruptions (especially power failures).

To begin, chores were formally assigned to strictly share the work on a rotating basis (Table 1). The actual time devoted to cooking and cleaning up varied between approximately 200 minutes and 350 minutes per day, with an average of 4 hours 23 minutes (Figure 1). The two people assigned to refueling the generator spent about 45 minutes a day (EOP and EOA, usually 3 times/day).

Table 1. MDRS5 Chore Assignments
(columns correspond to crew members; DGO = cook; EOP & EOA = engineering officer primary and secondary, for refueling generators)

	A	B	D	J	N	V
4/8/02		EOA		EOP		DGO
4/9/02			DGO	EOP		EOA
4/10/02		DGO		EOP		EOA
4/11/02	EOA			DGO		EOP
4/12/02	EOA				DGO	EOP
4/13/02	DGO & EOP		EOA			
4/14/02	EOP		EOA			DGO
4/15/02			DGO & EOP		EOA	
4/16/02		DGO	EOP		EOA	
4/17/02		EOA		DGO	EOP	
4/18/02		EOA			DGO & EOP	
4/19/02	DGO	EOP		EOA		
4/20/02		EOP		EOA		DGO
4/21/02				EOP		EOA

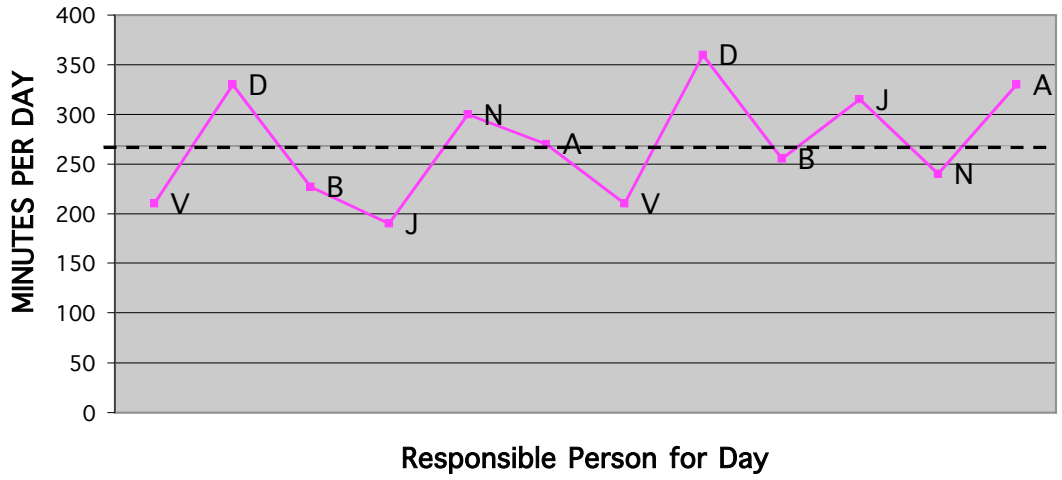


Figure 1. Total time devoted to cooking and cleaning by the assigned crew member, April 8-20; average 4 hrs 23 minutes.

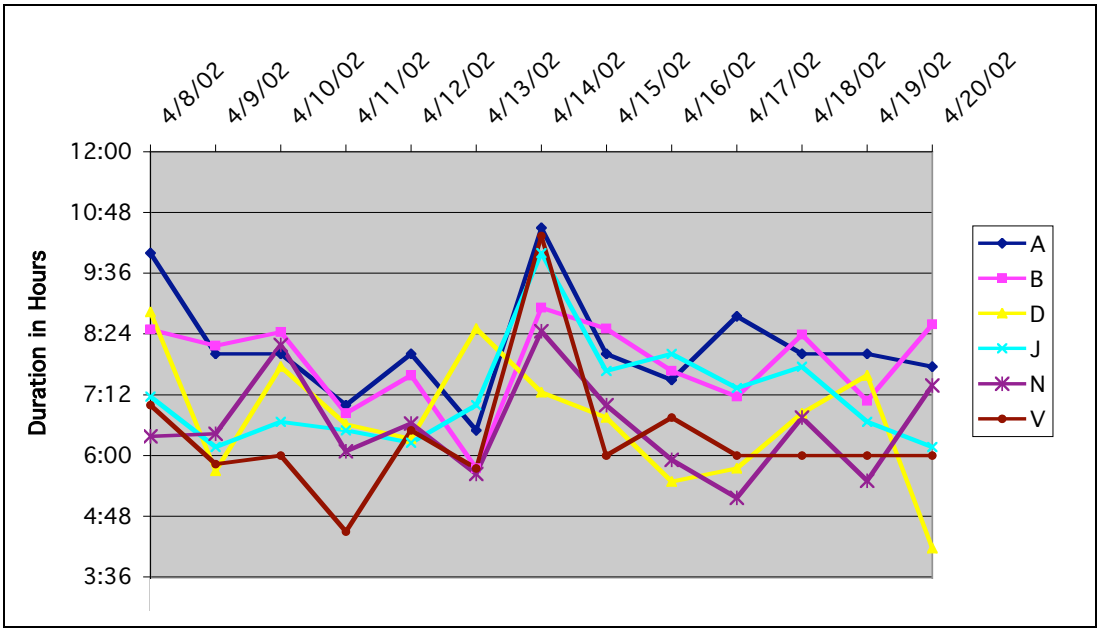


Figure 2. Sleep duration by crew members

Sleep time varied between 8:09 and 6:21 (h:mm, Figure 2). Duration is most strongly affected by the late evening movies and required attendance at the 9 am planning meeting. Note how sleep duration changes together for the group (4/14 was the Sunday without a meeting).

The extent of group activities including meals, meetings, movies, etc. can be determined within a few seconds accuracy from the time lapse record. The camera was placed on a tripod (Figure 3), using a wide-angle lens. As previously reported (Clancey 1999, 2000a, 2001a), the frames are examined manually, creating a spreadsheet of events and start/stop times, which is then processed by a computer program to produce tables for graphing. Table 2 summarizes the total time available to each person to work inside the habitat, taking into account their sleep, time devoted to chores, EVAs (and assisting others), lost time from power failures, movies, lunch, dinner, and meetings.

The big surprise here is the range of 2 hours between A and V because of V's shorter sleep and DGO time. These figures do not include irregular chores (helping the fill the water tank every few days), personal hygiene, or breakfast. The striking and important conclusion is that crew members do not have the same time available to do individual work. Note that N has the most time because she participated in fewer and shorter EVAs. A has almost fewer 1.5 hours available per day, which must be considered in evaluating her sense of frustration at lacking enough time to do her work.

Figure 3. Video camera set up for time lapse recording, with an example frame (Pletser is about to surprise the group with a box of candy)

Table 2. Time available to each person per day, after subtracting sleep time, group activities, and chores. Average is 10 hours 37 minutes.

	Average Available Time/Person/Day (h:mm)
A	9:15
B	10:11
D	10:57
J	10:35
N	11:30
V	11:16

To this point, we can relate individual differences and understand roughly how much time is available for laboratory analysis, writing reports, and so on. The initial question of "where does the time go" is now all the more mysterious—Why did the group report at the planning meeting on the fourth day (4/11) not having enough time, when everyone has at least nine unscheduled hours per day, and the average is 10.5? What the averages disguise is the change that occurred after this meeting (see Daily Schedule, Figure 4).

Creating the daily schedule chart was the most pivotal part of this analysis. It shows trends that were not visible or even known to members of the crew, such as the how having dinner earlier leaves open time for additional work before starting the movie. This graphic helps us understand at a glance how scheduling changes how much time is available when it is most needed, especially before and after dinner. The chart also makes us more aware that people do not operate on a 24 hour day (except for perhaps the journalist); every hour is not equally available for doing productive work. In practice, only the time between 9am and 10pm is universally available. Some people may rise early and work on e-mail before 9am and others may work well

past midnight. But if we are to understand how group scheduling affects productivity, we need to shift from studying individual averages to considering what time is *practically available* for everyone. Also, we need to examine the data in terms of variables that changed, looking for effects on productivity.

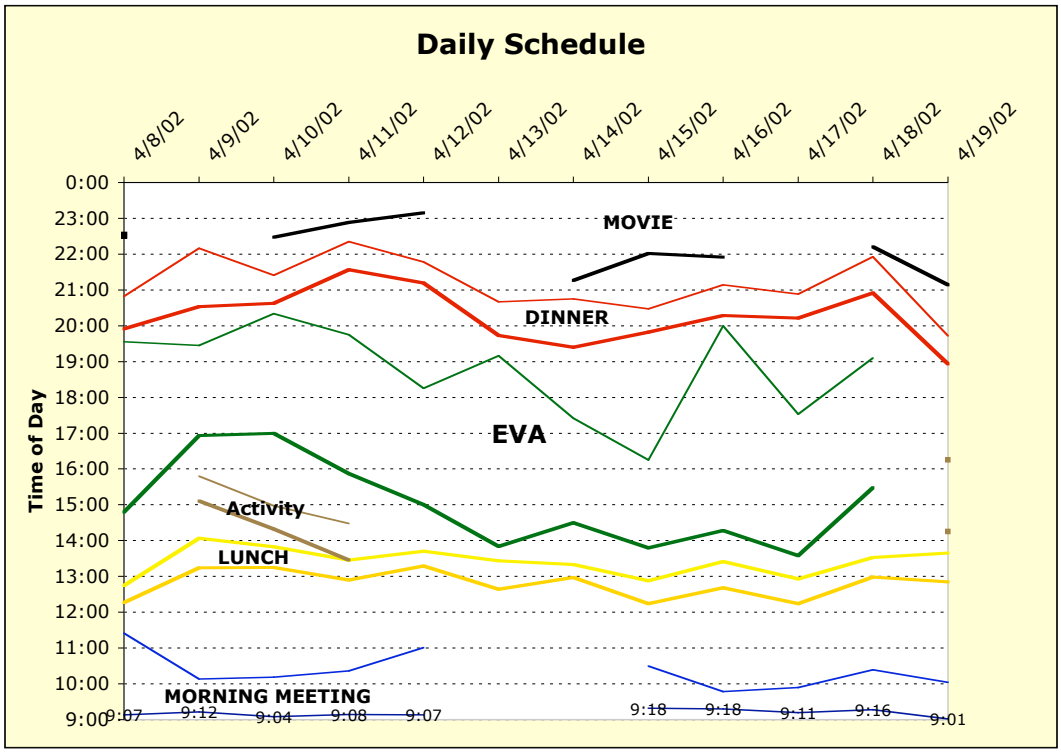


Figure 4. Extent of regular group activities (colors correspond to start and stop times of each activity, e.g., on 4/8 dinner began about 20:00 and ended about an hour later).

Figure 5 reorganizes the daily schedule data to show what time is equally available to everyone, comparing the two five-day periods previously described, omitting the rest, cleaning, and open house days. Productivity is increased by shortening the morning planning meeting and eliminating the afternoon tutorial activity. Again, we shift here from studying individuals to studying the resource available. Furthermore, productivity metrics can't be absolute; we must compare something. Given that the group discussed its productivity problems on the fourth day, a comparison of the first five days to the second five should reveal a significant cumulative change. In summary, our objective here is to *measure a valuable resource* that schedule changes (for example) might have affected.

What other changes might be useful? Lunch is already short (about 40 minutes); more might be lost in group cohesion by eliminating it as an organized activity. By making dinner earlier, more productive time is available in the evening. In the first week, the group worked as hard after dinner, but had to do this by making the movie later or skipping it. This effort probably cannot be sustained; one might further argue that 13 hour days are too long.

Notice that the shorter meeting time allowed an earlier lunch, plus it provided more work time (tempered somewhat by starting the meeting later). In turn, the EVA could be scheduled earlier (helped by improved weather that made waiting for the cooler part of the afternoon unnecessary).

Furthermore, the increase in productive time after EVAs is not at the expense of EVAs, which increased by 53% total person hours.

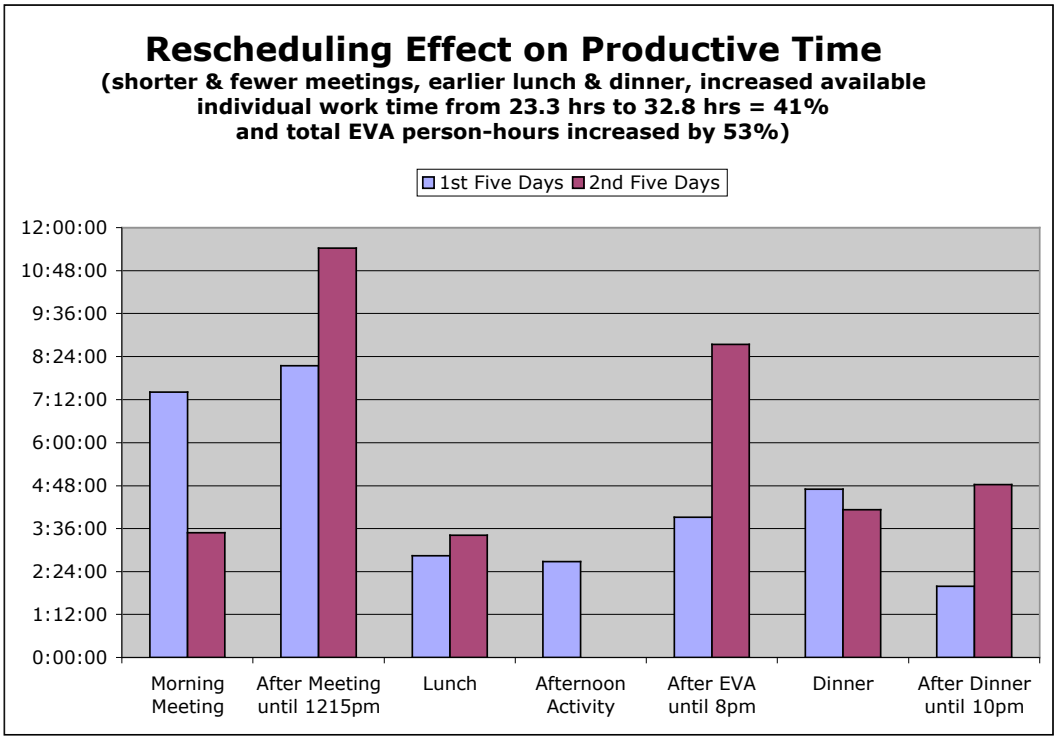


Figure 5: Affect of rescheduling on individual's available productive time (9am – 10pm).

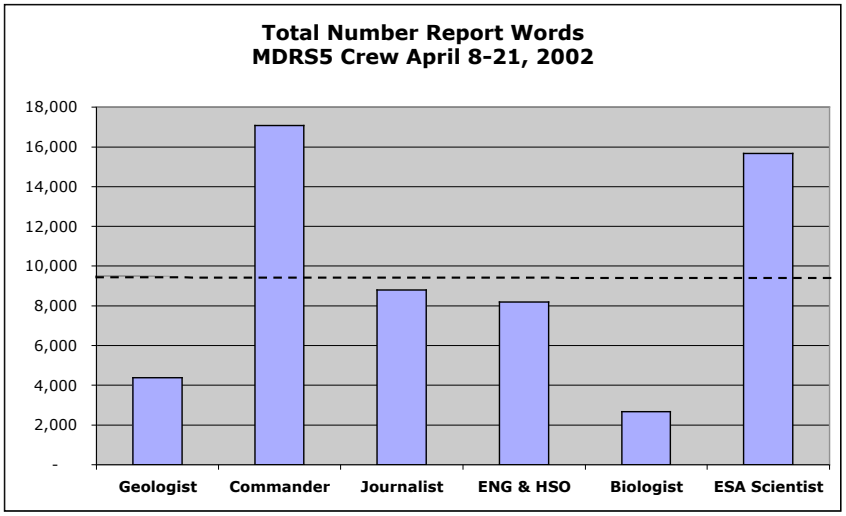
Productivity Metrics

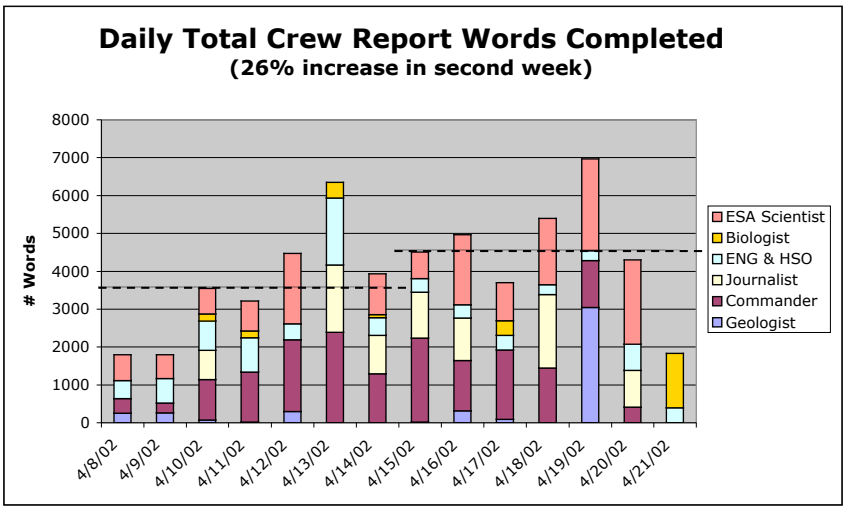
To this point we have considered only how much time the crew worked per day and how the schedule affected available individual time. Can we measure directly what the crew accomplished? I consider here the crew's reporting activity and to what extent daily plans were completed.

Report Writing

The crew wrote 97 reports over 12 days, totaling 57K words (Figure 6; for all MDRS reports and photographs, see www.marssociety.org/MDRS/2002Dispatches/). The commander, HSO, and ESA scientist wrote extensive daily logs (including French translations, not included in the total); the biologist and geologist wrote daily activity notes and weekly reports; and the journalist wrote five crew bios and two daily life stories. The number of words increased by 26% in the second week, with the most on the last Friday, suggesting an end-of-mission "completion effect" of submissions by the scientists and journalist (Figure 7).

The length of reports is perhaps the crudest metric of productivity. Nevertheless, reporting is an important part of surface operations work and the word count is objective. To be fair to the geologist and biologist in particular, other metrics must be considered: Number of samples taken, cultures grown, work in the greenhouse, the EVA area explored, number of waypoints found or logged, repairs, etc.





Figures 6 and 7. Total number of words in reports released for web publication by the MDRS5 crew and output per day compared over the two weeks.

Task Productivity

Experience in FMARS simulations (Clancey, 2000b, 2001b) suggested further study of how the crew plans daily activities. Each day the commander edited a table in a word processor to indicate what each person planned to do. Tasks were copied over from the previous day as necessary, with text changed to a strike-out font, to indicate lack of completion the previous day. Different formats were tried; the use of a simple table with one column per person and extra columns for group activities worked best (Table 3). These plans were analyzed by tallying the number of tasks proposed and deferred each day (Figure 8). On average two tasks were proposed per person/day; 60 were completed in the first six days, 72 in the second.

Table 3. Plan for Saturday, April 13 (strikeouts for previous day indicate tasks that were not started or abandoned; repetitions such as A's "EVA64 report" indicate continued work)

DATE	EVA	A	B	D	J	N	V	All
Friday 4/12/02 *Isolated T-storms 79*	*67 A, B, D* *Find Oyster route, photo collection* *Nancy's windsock and dry-site collection Pedestrian before 67*	*EOA* *Write EVA64 report*	*Must learn to use GPS, print map or take REI map* *EVA plan with A*	~~*Will interview V*~~ *"What are these people doing there all day?"*	*UPS debugging* *J will read Biolet manual*	*DGO* *Will write bio report* *Needs mostly dry sample from dust catchers*	*EOP* *Datalogger to Capcom & Gus* *Write GH report on Ecologger & watering*	~~*Fire drill with pump after mtg*~~
Saturday 4/13/02 Predict: Mostly sunny 81	68 J, V, D Mid-Ridge Planitia Loop (marking & finding waypoints)	EOP Update EVA calendar; Waypt ssheets; send 5 geo rpts; EVA64 map & coords; Primer	Review Andrew's & Frank's msgs; define an EVA "failure" experiment; Geo books pictures? Photos from loft Doubletalk	EOA Will interview V	Kodak help N; teach A how to create map from TopoUSA; Send audio file from 4/12 mtg	Sample grassy area? Set up new cols windy sample; Complete EVA 65 & Bio rpts. Learn Kodak	Reformat pics in Science report	**Latest day to refill water**

Note that April 14th is Sunday, the rest day. A strikingly even number tasks are proposed (15/day) with many additional tasks in the last week for wrapping up projects. Group tasks include EVA objectives, tutorials, and drills. Like counting report words, this measure is very crude—some tasks take a few minutes, others take days. Nevertheless, the comparisons between days should have some validity. Excluding the day of rest, whenever the group as a whole slept more than the previous night, the number of completed tasks increased compared to the day before. Otherwise, when the crew slept less, the number of completed tasks decreased or stayed the same.

Layout and Use of Space

Because no visitors were allowed inside the habitat during the 10 day closed simulation, we can completely characterize how space was used. The time lapse allows in principle determining, for example, when and for how long people used their staterooms. However, analyzing this data frame-by-frame is very time consuming and has not yet been attempted. Realizing this difficulty in advance, I instead recorded on paper where people were and what they were doing, every 15 minutes over a period of one and a half days (Figure 9). After this time, the patterns were obvious and it was clear that additional recording (at the 15 minute grain size) would provide little or no further information.

Direct observation showed that each person tended to spend most of their time in only a few (two to four) places. The three people with internet connections in their stateroom (B, D, V) spend most of the day there (A uses the stateroom to read before dinner; on other days J uses his upper bunk as a work table). Of the other three crew members, two (A & J) are working at the workstation bench of the upper deck (where they can connect their computers to the internet); on this day N is more often found in the laboratory of the lower deck. On this day D was the DGO, hence he is in the galley. The mess table, used for meetings, meals and video, is used evenly by everyone. ("Outside" refers to required work without suits; J worked with the water pump; B photographed an EVA activity.)

Comparison across two days shows fairly consistent use of space (Table 4). The second day EVA was longer and N worked at her workstation instead of the laboratory. Percentages are calculated by treating a person in a location as a "visit"; thus 6 visits are observed every 15 minutes; percentages given are based on number of visits in that location over the period indicated. Staterooms are treated as one place rather than six. Thus if everyone stayed in their staterooms for the entire observed period, it would be recorded as 100%. Similarly, 16% at the workstation on the first day could be one person for the entire period or some combination (see Figure 9).

Notice that about 25% of the day is spent sitting at the mess table. The lower deck is obviously underutilized; it is occupied only 8% of the time (though many visits to the toilet were not observed at the 15 minute grain size).

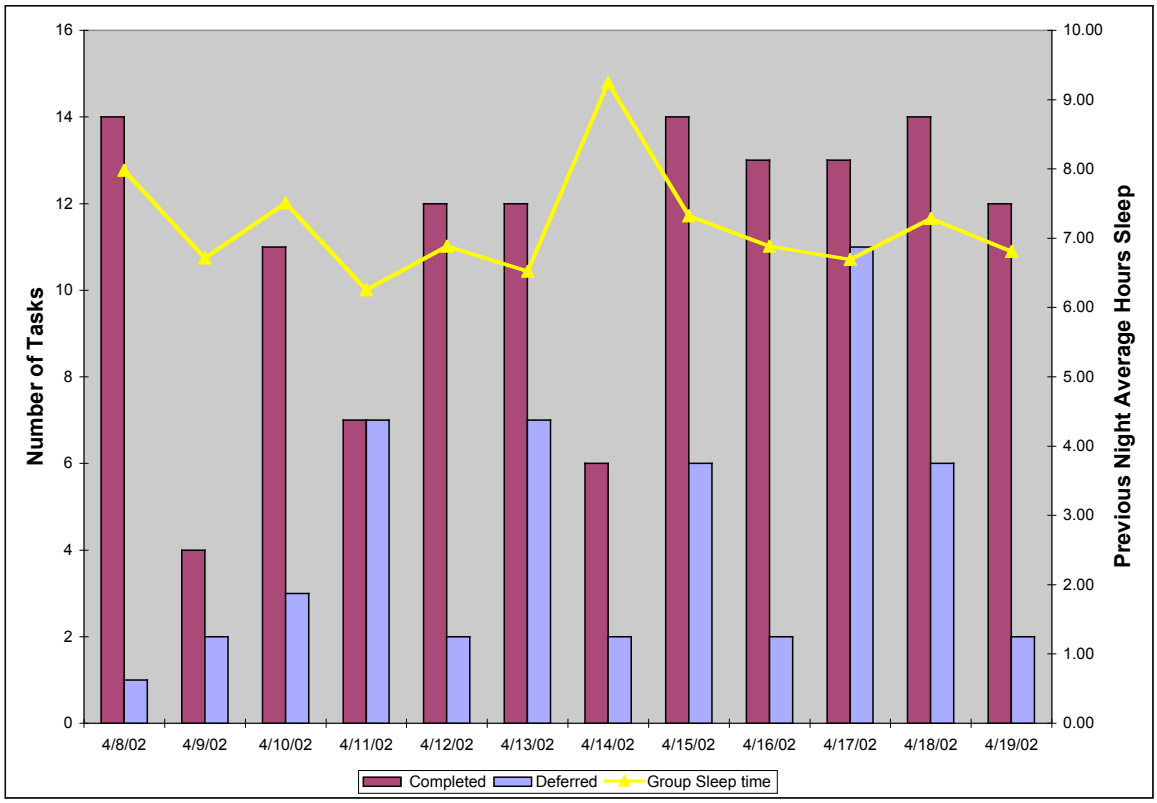


Figure 8. Tasks completed and deferred each day, plotted against average sleep time per person.

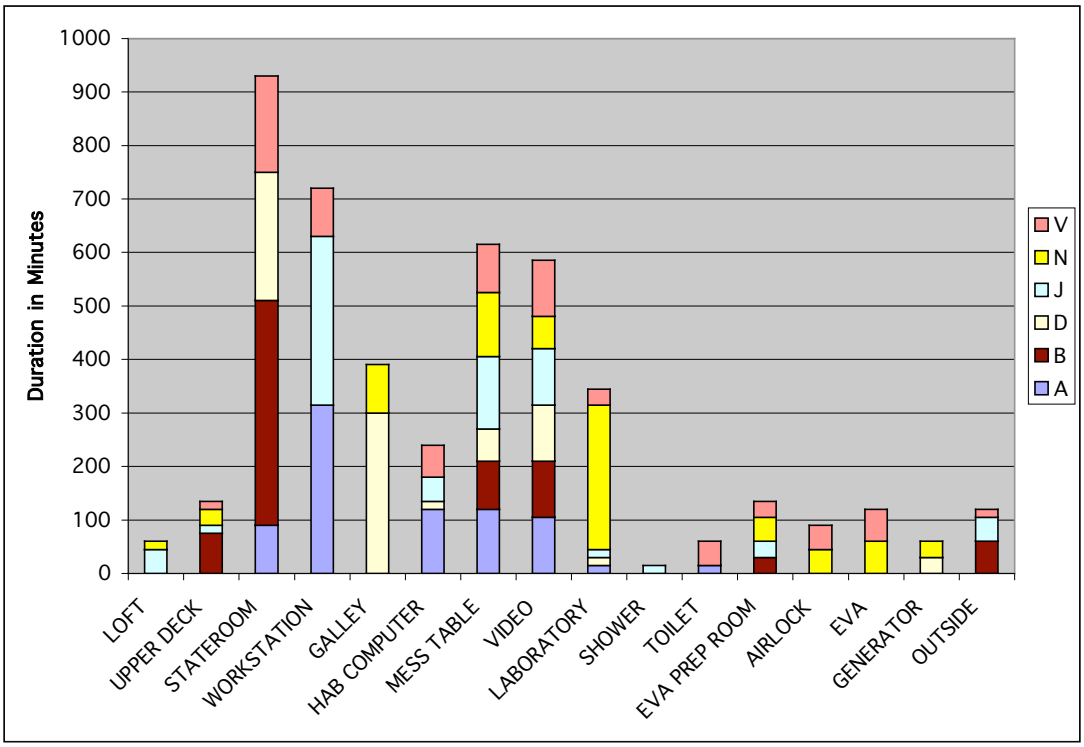


Figure 9. Total time in each location by person, MDRS5 11:15-24:00 April 15, 2002

Data about use of space can also be used to measure how often people move around. A "move" is defined as a person changing location from one 15 minute observation to the next. The first day has 52 observations; D moved the least, 17 times; V moved the most, 28 times; the average is 21. The second day has 69 observations; A moved the least (17); D moved the most (30); the average is 23. Further data are required to establish individual differences, if any. A person may also be very reactive—prone to get or check something when the idea occurs to him or her. Such short and frequent movements would not be caught by the 15 minute grain size.

Various commerical devices are now available for logging movements (e.g., used by retail stores for investigating where people look, what they touch, etc.). Although our time lapse video has this information, automated tracking is necessary for practically analyzing detailed movements over more than a few hours.

Table 4. Comparison of crew use of different habitat areas over two day. Occurrence >=5% shown in bold. See text for explanation of how percentages are calculated.

	11:15-24:00 April 15, 2002	7:00-24:00 April 16, 2002
LOFT	1%	0%
UPPER DECK	3%	3%
STATEROOM	**20%**	**20%**
WORKSTATION	**16%**	**12%**
GALLEY	**8%**	**5%**
HAB COMPUTER	**5%**	**7%**
MESS TABLE	**13%**	**16%**
VIDEO	**13%**	**9%**
LABORATORY	**7%**	1%
SHOWER	0%	1%
TOILET	1%	1%
EVA PREP ROOM	3%	**5%**
AIRLOCK	2%	0%
EVA	3%	**14%**
GENERATOR	1%	2%
OUTSIDE or GREENHOUSE	3%	2%

Crew Post-Occupation Survey

Crew members completed an individual written survey after the closed simulation ended (Appendix I). For the items ranked by importance, the most important (average rank >=4) were clean drinking water and sufficient power, closely followed by diet, adequate EVA suits, and toilet. Ranked next were free time, and showers.

From the open questions, the following patterns are apparent:

1. Everyone observed that the crew interacted harmoniously; our reactions to each other were without exception upbeat and cheerful.
2. Everyone reported either insufficient time to accomplish objectives and/or inadequate leisure time: Computer network problems and/or interruptions are cited by everyone, but only two people mention unnecessary report writing (in such a simulation, the audience can be ill-defined).
3. Four out of six (4/6) said the most important problem is the toilet facility (followed by power).
4. Providing stable, sufficient power, without refueling would have the greatest effect on morale and productivity. Power outages halted all activity (loss of about an hour per day) and exacerbated the computer network difficulties (only three internet lines, tenuous

connections, and low bandwidth). (A new diesel generator was later installed with greater power output and less maintenance rquired.)

5. Tool usage posed problems cited by everyone, either from lack of access to a tool (computer or camera), lack of knowledge to use a tool, or being interrupted by requests for assistance to use a tool.
6. The best moments were outdoors (5/6), with three people mentioning a particular EVA in which they participated together (this is remarkable given that surveys were privately prepared).
7. The worst moments were all different and had no resemblance to each other (e.g., as commander, my worst moment was when a crew member felt ill during dinner; nobody else mentioned this event, not even this crew member).
8. For those having a computer connection, the stateroom was an important place, for others it was just a place to sleep; half mentioned the importance of privacy.
9. Everyone wanted a better EVA suit and more than one shower/week.
10. Everyone would have continued a third week (though one person couldn't get time off from work); everyone would stay for a month if family and work allowed.
11. Food and habitat temperature were non-issues for this rotation.
12. The group provided varied and imaginative ideas for habitat improvement, with a surprising number of suggestions focusing on kitchen equipment.

In response to the question of "where did the time go," every crew member would try to change how they used their time on subsequent simulations:

- Geologist: wanted to find a way to work uninterrupted for longer periods of time; couldn't concentrate on reading and writing.
- Commander: never read any of the crew's reports (!), too much time writing logs.
- Journalist: didn't write enough "daily life" stories; on reflection would have done this instead of writing biographies.
- Health and Safety Officer : too much time being "Mr. Fix-it"; didn't have time to create a systems model of the habitat as planned.
- Biologist: insufficient lab time, too much group time (ironically, this person had the most amount of available time per day, considering sleep, EVAs, and chores).
- EVA Scientist: insufficient EVA time, too short and lacking adventure.

Individual differences in what individuals didn't say are also intriguing:

- Only one person (the commander) didn't complain about lack of time (rather I wanted entire days off, to do something entirely different).
- Only one person (the journalist) didn't mention outdoor activities as a highlight.
- Only one person (the biologist) didn't complain about interruptions or network problems, perhaps because she spent so much time working alone in the laboratory, away from her computer.

- Only one person (the EVA scientist) didn't complain about biolet and mentioned the importance of skill training in advance (reflecting his professional experience).

These differences may reflect mood or experience on the day of the survey. In any event, we are reminded that individuals will react differently to identical circumstances, sometimes because of different roles, but also because of temperament and past experience. The comments highlight that differences must always be expected, even in a crew that experiences itself as being especially harmonious. Further, over a long mission, the crew might enjoy varying its routine, so for example, one person could be left alone to work and eat independently for a day or a week, according to his or her personal need for concentrated effort and variety.

DISCUSSION: SYSTEMATIC WORK SYSTEM DESIGN

In this section I summarize the framework of analysis and modeling that orients the empirical study carried out in MDRS5 and comment on how the analysis is informing ongoing research on work systems design (Clancey, et al. 1998; Clancey, 2001b, in press; Sierhuis, 2001). In this context, the *work system* (Figure 10) refers to the design of the MDRS habitat and its tools, the crew's roles and assignments, the daily schedules, protocols for interacting with mission support, and practices for using the habitat, tools, and interacting with each other.

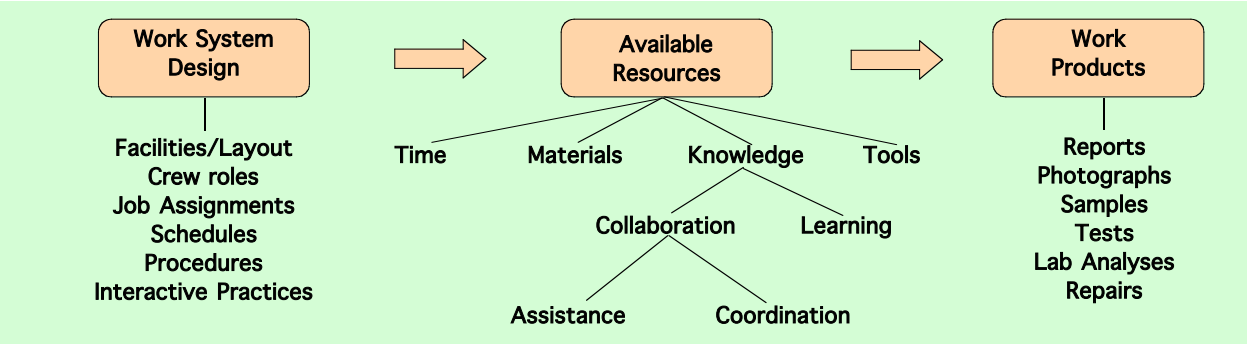


Figure 10. Different work system designs (facilities, tools, crew roles, etc.) affect the resources available, which affects the quality and quantity of work products.

The framework suggests that, in order to understand how a given work system design causally influences productivity (the quality and quantity of work products), work system studies should focus on *changes in resources* brought about by schedules, facilities, roles, processes, etc. For example, this paper has shown how the crew's change in the daily schedule, prompted by a desire to be more productive, resulting in a significant increase in the amount of time available for individual work, such as report writing. The next step is to show how a change in resources actually affected work products (e.g., number samples processed).

In previous reports about the Haughton-Mars Project and FMARS (Clancey 2001b, in press), I have mentioned a simulation method for representing work practices, using a tool called Brahms (Clancey, et al. 1998; Sierhuis, 2001). The question arises whether a Brahms simulation of MDRS, which is so tedious and expensive to construct, would add anything to a spreadsheet analysis. On the one hand, the analysis given here reveals what a generic habitat model must

contain, and how it must be parameterized (the work system design elements). For example, the DGO schedule would be represented as a document, posted on a whiteboard near the mess table. The schedule is uniformly visible where it can be referenced by individuals and during meetings.

The most obvious use of a Brahms simulation would be for what-if analyses, in which we would experimentally varying the work system design. For example, we could have the simulated crew follow different daily schedules, to show how this influences productivity. Or we could have the crew adapt different protocols for (not) interrupting each other. A simulation is useful when many variables interact and have longer-term consequences. For example, suppose that the crew were not allowed to interrupt each other during the morning work session, between the planning meeting and lunch. Individuals needing assistance may have their work delayed, affecting the entire crew later in the day. Or on the contrary, the crew may learn to depend more on mission support, which even with the time delay may be a better resource than other busy crew members.

What if one person were always the DGO or the generator were always handled by one person? One scientist thought such changes would improve her productivity. However, ethnographic studies of scientific expeditions organized in this manner (Bernard and Killworth, 1974) show that severe interpersonal difficulties may occur, as all must live together in an hazardous, isolated environment, but some are reduced in status to being servants to the others. But other changes are possible: What if the DGO assignment could be freely traded by crew members, so they could work on different days or do something else in exchange or be freed of the task, in recognition of another contribution to the mission?

Based on the analysis of MDRS5, the most pressing use of a Brahms simulation model would be to help us formulate and understand how interactions in crew activities and their individual problems cause interruptions for others and thus reduce productive time (including time to get back to work). However, creating such a model would require observations that were not made during MDRS5. Thus, we see the familiar pattern that a study reveals a topic of interest, and we must return to the field with different questions and additional observational methods. In particular, individuals must be tracked more closely during the day, with a better understanding of how they organize time, in order to understand the cause of interruptions and their effect.

LESSONS LEARNED ABOUT ANALOG RESEARCH

Of paramount importance, given the effort and expense to build a research station like MDRS, is determining what can be learned from an analog activity and how the activity should be managed and controlled as a scientific investigation. The key findings from MDRS5 are:

1. **Adjusting the group activity schedule creates more *useful time* for working** (before lunch, before and after dinner).
2. **Interruptions significantly affect productivity:** power failure, group activities, assigned chores, requests for assistance, computer network problems, incoming email.
3. **Timing and counting activities (systematic recording) is essential for detecting patterns and making work system design recommendations**. The process is iterative; a successful outcome is determining what observations need to be made on subsequent controlled simulations.

From a work systems design perspective, the chief finding is that rescheduling group activities can increase the amount of available, useful time for individual work. Specifically, we shortened the morning meeting (most likely because less planning was necessary), we had earlier (but not shorter) meals, and we eliminated the after lunch "tutorial." In this analysis, time after 10 pm is not deemed as available (not useful for getting work done). An important consequence of shifting the day downwards was that movies could begin earlier (by 10 pm) and people could get to bed sooner (by midnight). This lowered fatigue on the subsequent day and increased the number of proposed tasks at the planning meeting.

Secondarily, observations (supported by the post-rotation crew survey) showed that interruptions negatively influence productivity. These include: power outages, group activities, attending to assigned chores, requests for assistance, computer network problems, correspondence by email, lack of access or knowledge how to use tools (such as digital cameras and GPS).

Methodologically, the analysis demonstrates that systematic recording enabling timing and counting activities (e.g., time lapse) is essential. In particular, the above findings are based on comparative analysis that measures total group activities, resources, and products in different time periods, before and after a change in the work system. The analysis requires identifying resources that affect work product quality and quantity and that are affected by the work system design (WSD). Whereas measures of productivity may be always incomplete and problematic (e.g., number of words written, number of photographs, duration and extent of EVAs), one can often measure intermediate factors that are universally available for different purposes (e.g., useful time available, internet bandwidth, number of chairs) and then show how these influence productivity. Nevertheless, although the crew reported lack of time as the chief frustration, one solution is to redesign the reporting products required and hence the schedule. Thus it is essential to keep in mind the relationship:

WSD (e.g., schedule) -> RESOURCES (e.g., time) -> PRODUCTS (e.g., reports)

Changing the WSD affects resources, but one may also redefine the outputs required.

Finally, the study showed the importance of knowing what you want to measure in advance, so recordings are taken systematically. Nevertheless, this understanding develops and changes as the study proceeds. For example, the experience in FMARS2001 suggested that full time lapse over the entire rotation was desirable. Then during MDRS5 I discovered that following individuals to understand how they were interrupted was also necessary.

Lessons Learned about Metrics

This study has considered different metrics for describing the work system and productivity. Given the well-known importance of metrics for engineering, it is useful to summarize how the concept carries over to work system design:

- Metrics, as considered here, require *counting*, which requires *systematic observation over some domain of analysis*.
- A *domain of analysis* consists of a time period and some combination of one or more of the following: places, kinds of events, individuals (or a group), artifacts (e.g., a tool), and information (e.g., a task in a plan). For example, I measured how often people moved during 15 minute time periods; I measured how many tasks were proposed by the group

each day; I measured the duration of power outages. I could use the time lapse to count other things, such as how long people used the habitat computer and how many times people visited their stateroom during the day.

- An observer must avoid two extremes: 1) observation focused only on predefined metrics (in order to be precise) which produces no surprising kinds of observations (e.g., the journalist noticed how the chairs were repeatedly moved to five different areas), and 2) observation so general and unsystematic (in order to be unbiased) it produces no measurements at all (e.g., my first-timer's experience during HMP-98).
- Metrics of value will not all be anticipated; you will discover what can be counted in the data if you have cast a broad enough net (e.g., the EVA reports enabled observing that every possible group of two people shared at least one EVA). You cannot know for sure in advance what counts will produce meaningful metrics.
- Work system design metrics are *multidimensional*: they always include a time dimension, and they are comparative (planned vs. actual, one time period vs. another, one individual vs. another).
- Cumulative group productivity metrics may be essential where different roles would make averages misleading or irrelevant (e.g., number of person/hours devoted to EVAs in a given time period).
- Data often needs to partitioned to reveal key variables (e.g., the logic of viewing the rotation as two 5-day periods rather than 12 uniform days). An overview trend graph (e.g.,, Figure 4) may show a gradual change as a new policy or strategy is adopted, thus suggesting a "before and after" partition.
- Resources should be identified and measured, especially when products are difficult to quantify.

Lessons Learned about Photographic Observation

The following are some lessons learned about using photography to record human behavior:

- Regular observations are more amenable to analysis than raw video or unsystematic photographs. Decide in advance some category you want to photo-document (e.g., the phases in putting on a suit, all uses of duct tape, use of staterooms). Create a list and refer to it during the simulation.
- Common photographic mistakes:
    - Focusing on individuals and not photographing the context (better to have a wide angle camera fixed on an entire meeting table, than to keep focusing on individuals).
    - Clear, audible sound is more important than close-up images.
    - Forgetting to set and synchronize clocks in recording devices.
    - Not bringing enough film, tape, or storage (in the guise of reducing costs).

- Observational heuristic: When two or more people are together, photograph the setting, the details of tools and documents, people pointing, and people using tools. Ask yourself what they are doing and document how the interaction develops.
- Time lapse is especially good for capturing the start and end (and hence duration) of group events (especially meetings and eating). When placed appropriately, it can show when people enter and leave different spaces (e.g., staterooms).
- Don't worry about watching, let alone analyzing all your video—the trick is to capture the moments that need to be studied. For example, by recording every planning meeting, I captured the discussion of the fourth day when productivity problems led to schedule changes, and this was one hour well worth studying.

Lessons Learned about Analyzing Behavioral Data

The following are lessons learned using spreadsheets and charts to analyze behavior data:

- Prepare spreadsheets to organize the data and then charts.
- Use macros (e.g., programmed using the Visual Basic language built into Microsoft Excel) to count occurrences, producing additional tables and charts (e.g., counting how many times people moved from the original spreadsheet that recorded where people were every 15 minutes).
- Create many kinds of charts to discover meaningful trends or comparisons. For example, using Microsoft Excel's Chart Gallery preview option, explore what kinds of graphics produce readable (and hence potentially useful) displays.
- Find ways to meaningfully partition data by grouping people, places, and days in different ways.
- Categorize periods by abstracting how time is used: "morning individual work time" "pre & post sleep" "productive work time" (9 am – 10 pm).
- Compare periods to understand schedule changes, periodicity, and carryover effects.
- Identify recurrent events not recorded systematically (e.g., interruptions).
- Look for invisible patterns and practices (e.g., frequency of movement).

CONCLUSIONS AND NEXT STEPS

The fifth rotation of MDRS met all of its primary objectives—in field science, in reporting, and in the study of human exploration. This rotation introduced a number of creative innovations to the Mars Society's analog program: the closed simulation protocol, systematic recording of the entire rotation, a multiple failure IVA/EVA simulation, a plant growing psychosocial experiment, a resident journalist, an international press open house day, and many improvements to the habitat (bread maker, video/slide screen, medical inventory, lab tools inventory, computer connections document). The group generated 97 reports, including a geology primer for the region and comprehensive reports on the greenhouse datalogger and soil cultures of the surrounding environment. The data recording methods were entirely successful, allowing discovery of unexpected trends in group behavior, showing how schedules influence

productivity. The analysis was related to ongoing work in work practice simulation to show how statistical data informs modeling, while revealing questions for further observation.

The following are ideas for future work:

- Use of staterooms: Use the available time lapse to determine the number of visits and duration per person.
- Experiment with an internet connection in all staterooms: Does everyone choose to work inside and ignore the outside workstation area? Or do one or two people adopt the external area as a personal desk?
- Develop tools for automating EVA navigation and science data logging (Clancey, et al., in press).
- Study interruptions by following ("shadowing") individuals for at least three hour periods. Record the causes, durations, and any apparent disruption in other activity.
- Ask individuals to track time spent on activities of importance to them (i.e., let them decide what to record).
- Provide verbal EVA reports and let mission support write the formal reports
- Compare use of public and private space in MDRS and the EuroMars habitat.

Although the closed simulation was beneficial for carrying out a systematic study of use of time and space in the habitat, this is by no means the only scientific way to use MDRS or other analog research stations. In particular, one can carry out controlled protocols for shorter periods of time, say a few hours or a day, in which experimental equipment and procedures are used. MDRS's setting makes it especially attractive for research on EVAs, including different configurations of suits, rovers, robotic assistants, agent-based software, local capcom monitoring, and remote mission support.

Finally, it is worth reviewing the scope and dimensions of the observations and analysis I have carried out in the HMP, FMARS, and MDRS settings since 1998:

- The organization and activities of a field science expedition:
  - Multiple phases/rotations with handovers
  - Multiple field seasons, showing development of exploration
- Metrics of work practice based on the rotation, subgroups, individuals
- Individual scientific discovery and reporting genre
- Ensemble behaviors (micro or interaction analysis; e.g., process of starting and ending meetings)

My impression is that carrying this work forward will benefit from focusing more on individual's experiences as they plan, learn, and cope with the broader expedition setting.

ACKNOWLEDGMENTS

This work was made possible through the vision, volunteer work, and financial contributions of the Mars Society and its supporters, especially Robert and Maggie Zubrin, Frank Schubert, and Tony Muscatello. Mission support was provided by Mark Klosowski, Frank Crossman, and other members of the Northern California Mars Society. Discussions with Boris Brodsky, the Brahms modeler for FMARS and MDRS, have been very helpful. Related work pursued by Maarten Sierhuis and Alessandro Acquisiti, in simulating how ISS crew members follow schedules, has inspired many aspects of MDRS5 observation and analysis. The crew members provided essential information for carrying out this study, as well as helpful comments for improving the presentation. 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 and the University of West Florida.

REFERENCES

Bernard, H. R. and Killworth, P. D. 1974. Scientists and crew: A case study in communications at sea. *Maritime Studies and Management,* **2**: 112-125.

Clancey, W. J., Sachs, P., Sierhuis, M., and van Hoof, R. 1998. Brahms: Simulating practice for work systems design. *International Journal of Human-Computer Studies*, **49**: 831-865.

Clancey, W. J. 1999. *Human Exploration Ethnography of the Haughton-Mars Project 1989-1999*. Presented at the Mars Society Annual Meeting. Boulder, CO. In R. Zubrin and F. Crossman, *On to Mars: Colonizing a new world,* Burlington, Ontario, Canada: Apogee Books, published 2002, CD-ROM.

Clancey, W. J. 2000a. Visualizing practical knowledge: The Haughton-Mars Project. (Das Haughton-Mars-Projekt der NASA — Ein Beispiel fur die Visualiserung Praktischen Wissens). In Christa Maar, Ernst Pöppel and Hans Ulrich Obrist (Eds.), *Weltwissen - Wissenswelt. Das globale Netz von Text und Bild*, pp. 325-341. Cologne: Dumont Verlag.

Clancey, W. J. 2000b. A framework for analog studies of Mars surface operations: Using the Flashline Mars Arctic Research Station. Presented at the Mars Society Annual Meeting. Toronto. In R. Zubrin and F. Crossman, *On to Mars: Colonizing a new world,* Burlington, Ontario, Canada: Apogee Books, published 2002, CD-ROM.

Clancey, W. J. 2001a. Field science ethnography: Methods for systematic observation on an Arctic expedition. *Field Methods*, **13**(3), 223-243, August.

Clancey, W. J. 2001b. Simulating "Mars on Earth": a report from FMARS Phase 2. Presented at the Mars Society Annual Meeting. Stanford, CA. In R. Zubrin and F. Crossman, *On to Mars: Colonizing a new world,* Burlington, Ontario, Canada: Apogee Books, published 2002, CD-ROM.

Clancey, W. J. (in press). Simulating activities: Relating motives, deliberation, and attentive coordination. *Cognitive Systems Research*, special issue on situated cognition.

Clancey, W. J., Sierhuis, M., Kaskiris, C., van Hoof, R. (in press). Brahms mobile agents: architecture and field tests. *Proceedings of the AAAI Symposium on Human-Robot Interaction*. North Falmouth, MA. November 2002.

Perrow, C. 1999. *Normal accidents: Living with high-risk technologies.* Princeton: Princeton University Press.

Sierhuis, M. 2001. *Modeling and simulating work practice*. Ph.D. thesis, Social Science and Informatics (SWI), University of Amsterdam, SIKS Dissertation Series No. 2001-10, Amsterdam, The Netherlands, ISBN 90-6464-849-2.

Stuster, J. 1996. *Bold endeavors: Lessons from polar and space exploration*. Annapolis: Naval Institute Press.

APPENDIX I. MDRS5 Crew Survey

1. Did you go anywhere to be alone or to seek quiet?
2. What is your overall impression of your stateroom in a word or phrase?
3. Did you use your stateroom during the day? For what purpose?
4. Was there sufficient work space?
5. Did you ever work on the lower deck? For what purpose?
6. Were you productive? In what way or why not?
7. What was your most memorable positive moment (or moments)?
8. What was your most memorable negative moment (or moments)?
9. What is your assessment of the crew selection in terms of quality of knowledge, skills, and temperament?
10. Did you feel you had enough time for your research, i.e., sufficient field and lab time?
11. What reports did you write? Was this too much or less than you desired?
12. Would you be happy if you were staying another week?
13. Would you participate in a longer simulation of 4 weeks? 8 weeks? 12 weeks?
14. Did you have sufficient electricity? Water? Food?
15. Was the hab temperature comfortable? Was the ventilation adequate?
16. Was the amount of unscheduled time appropriate given the group's goals? Sufficient from a personal perspective?
17. Any suggestions for improving the EVA suits and equipment?
18. Did you like the evening group activities, or would you prefer to use the time for personal reading, working, etc.?
19. Did wearing the suits hinder exploring this area? In what way?
20. What is the one thing you would most like to see improved in the hab?

Please rate the following for their importance to you on a scale of 1 to 5, 1 being Not Important and 5 being Very Important

1. Clean drinking water --
2. Sufficient and reliable electrical power --
3. Unscheduled time for sleep and unscheduled activities --
4. Reliable and effective toilet facilities --
5. Capable and simulation realistic EVA suits and equipment --
6. Wholesome and varied diet --
7. Work areas with adjacent space for personal items (e.g., notes, drinks) --
8. Entertainment (e.g., DVD movies) --
9. Quiet and comfortable private space –

10. Cleanliness (showers, hot water) --
11. Personal storage areas --

APPENDIX II. Simulation of a Normal Accident

At the midpoint of the closed simulation, the crew had one relatively unstructured day. During the afternoon, a previously agreed upon script was enacted as an experiment on time-delayed communications and emergency management. The scenario was deliberately designed to avoid serious concern by mission support, meaning that none of the events would suggest that escalation was required (e.g., notifying Mars Society headquarters). Nevertheless, the events were designed to be realistic and dramatic.

The scenario was inspired by a framework provided in *Normal Accidents* (Perrow, 1999), such that multiple problems occurred with different kinds of causes: becoming lost during an EVA (*human error*), a stuck zipper (*mechanical failure*), wind and heat (*environment condition*), and radio problems (*system design*). According to Perrow, accidents in complex systems cannot be completely eliminated. Rather they must be viewed as part of normal operation, however infrequently they may occur. The design of a work system must accordingly take into account tools, procedures, training for handling emergency, critical situations. The purpose of our experiment was to first develop a method for simulating multiple failures in the MDRS setting and second to understand how multiple failures would be coordinated by the commander and mission support.

The method for this simulation included a detailed script, indicating what individuals would do and what they would report at five minute intervals (Table II.1). Special for this experiment, audio recordings (using the Radio Shack DSSPlayer) of communications between remote teams and the habitat capcom were transmitted to mission support by email with an average five-minute delay (roughly the time required to download the file from the recorder to the computer, to compose an email, and for the message to be received). Additionally, mission support responded with a simulated five minute time delay. Coordination and communication occurred using handheld walkie-talkies. The capcom would prompt crew members according to the script to remind them of their lines. For each recording, capcom would then say, "Okay, record now," and the crew member would call in using the usual protocol (e.g., "Capcom, this is Nancy. It's really windy out here…"). The script took two hours to enact, as designed. Mission support suspected something unusual was occurring, but for the most part viewed the scenario as a a real series of events.

The plot: Capcom is on the MDRS upper deck monitoring a simultaneous IVA (D and V working inside the greenhouse) and an EVA (N, J, A walking to a site about 100 meters south of the habitat). The beginning appears normal. After a radio check, N, J, and A suit up, egress, and proceed to a previously visited location to deploy and retrieve sample containers. Meanwhile, D and V have entered the greenhouse to do something with the datalogger. The problems begin as indicated in the script (Table II.1), with new important information being received by capcom every five minutes. The script is such that with the time delay, any sense of priority or problem resolution plan must be reconsidered just as it would be transmitted by mission support. Even with the minimal time delay (five minutes), mission support is unable to stay on top of a fast-changing situation, one that requires revising plans as new information comes in. At best they can follow along. Capcom, too, is overloaded. At best he can put the crew on "safe hold" and

focus on the most important problems. For mission support, the situation changes too quickly to provide any useful advice. In a real Mars emergency, the time delay is likely to be longer and the information available perhaps less complete, in particular, automated methods would be required to keep mission support aware of what is occurring, as capcom will be too busy interacting with the crew and local systems to be sending and receiving email. Even within this scenario, capcom often forgot to check whether mission support had sent a message.

Table II.1 Script excerpt for multiple failure accident

Time (mins since start)	Actor	Action	Actual time
0	N	Reports she is alone	15:45
5	B	Reports no comms with J &A	15:50
10	D&V	Report zipper on Greenhouse door is stuck, so they are trapped inside	15:55
15	N	Reports she is very hot and walking back to the hab (but goes via waypoint 102 to fetch camera)	16:00
20	B	Reports via email that UPS is beeping, and he must refuel in 30 minutes	16:05
25	D&V	Ask for permission to cut or rip open the door	16:10
30	J&A	Report they are back at potholes, but N is not in sight	16:16
		(N has returned by way of the wind catcher hill so she can follow the main road)	
35	D&V	Are arguing about what to do	16:20-16:24

APPENDIX III. Suggested Research Station Improvements

The following are some of the crew's suggestions for improving MDRS or similar habitats:

- Recognize that the EVA prep room can serve as a private meeting area; provide chairs that stay there.
- List available tools and movies in the manual (so future crews know what to bring)
- Show new crew members how an upper bunk can be used as a workbench; provide a stool (perhaps opening from the wall) to enable using that area instead of the built-in desk.
- Provide the following in every stateroom: a mattress, a porthole, an internet connection, book shelves, more clothes hooks.
- Lower deck is only used for work by two people, an inefficient use of open space. Consider how it might be outfitted to draw the group downstairs for variety (e.g., MDRS7 used it as a music studio).

- Provide better kitchen appliances for more efficient use of crew time. The crew made extensive use of the crock pot (slow cooker) and bread maker.