Evaluation of Human vs. Teleoperated Robotic Performance in Field Geology Tasks at a Mars Analog Site

B. Glass, G. Briggs, J. Jasper[†] and K. Snook[*]
NASA Ames Research Center
Moffett Field, CA 94035, USA
brian.j.glass@nasa.gov

Keywords:
Planetary exploration, rovers, field tests, human exploration, robotics, science return, space.

Abstract
Exploration mission designers and planners have costing models used to assess the affordability of given missions – but very little data exists on the relative science return produced by different ways of exploring a given region. Performing cost-benefit analyses for future missions requires a way to compare the relative field science productivity of spacesuited humans vs. a virtual presence/ teleoperated robot or rover from a nearby habitat or orbital station, vs. traditional terrestrial-controlled rover operations. The goal of this study was to define science-return metrics for comparing human and robotic fieldwork, and then obtain quantifiable science-return performance comparisons between teleoperated rovers and spacesuited humans. Test runs with a simulated 2015-class rover and with spacesuited geologists were conducted at Haughton Crater in the Canadian Arctic in July 2002. Early results imply that humans will be 1-2 orders of magnitude more productive per unit time in exploration than future terrestrially-controlled robots.

Introduction

Separate, often competing camps of opinion regarding planetary exploration have fostered humans-as-explorers vs. cheaper robotic alternatives. Both sides have made credible arguments regarding the likelihood of mission success, science return, adaptivity, and relative costs. While missions continue to be proposed with either humans alone or robots alone, these competing approaches allow no synergy between the human and robotic programs. [1]

Robotic surface exploration can add a broader set of sensory inputs and observe more of the spectrum than can human eyesight [2], helpful in mineral and chemical identification and other scientific measurements. Safety and life support issues are reduced or eliminated, costs greatly reduced and improved strength and endurance are possible. These arguments have led some authors to question sending humans into space at all.

However, even with the expected advances in robotics over the next decade or two, robots will still lag human capabilities in real-time perception, planning and recovery from, or adaptation to unexpected or adverse circumstances [3]. Increasing levels of detail in geological or biological field work require real-time decisions on which subunits to choose to measure or sample, without a priori knowledge [4]. Detailed field work may then require human cognitive and perceptual capabilities.

A recent NASA study [3] surveyed the robotics community to assess the likely and possible capabilities of space robotics over the next 10-15 years. It found that even optimistic estimates of future overall robotic capabilities would not approach those of humans in exploration.

One difficulty in comparing humans vs. robots in space exploration has been that cost-benefit trades have not been possible. Cost estimates for various mission scenarios are readily available, but there has been a lack of metrics for assessing exploration benefits. What defines science return, or productivity? Humans probably make better geologists than robots [5, 6], but by how much?

This paper will describe an initial attempt to measure the science productivity of each exploration approach by comparing the field observations returned by humans in simulated extra-vehicular activity (EVA) vs. simulated future robots controlled by a remote science team. After defining figures of merit, the paper will discuss the expected capabilities of 2015-era surface exploration robots. The field experiment will then be described, including controls, siting, and initial conditions. The paper will conclude with the initial results and conclusions deriving from the summer 2002 field season's experiment.

[†]QSS, Inc. ; [*] Now at NASA-Johnson Space Center

Formulation Issues

Definitions of science productivity in exploration

What constitutes "science return", when a single given dataset or sample may (historically) prove to be the key to understanding a site or confirming a given theory? Bulk digital storage was considered, for instance, but discarded as much returned data is unproductive -- redundant, devoid of content or noisy. A set of several possible metrics for assessing field exploration were compiled, after interviewing current practicing field geologists and planetary scientists. The initial set, as shown in Table 1, became a starting point for performance discussions.

Table 1. Proposed metrics for human/robotic traverses.

Metric	How Measured
A. Total distance	GPS waypoints recorded onsite by test director; video footage
B. # stops	Video footage, notebook entries, # of rover panoramas
C. # sites visited	As for stops, minus revisits (alternate metric)
D. # of samples kept	Counted post-test
E. # of image-retakes	# of closeup image files
F. # Site descriptive phrases	Text analysis of descriptive report
G. # hypotheses	Text analysis of descriptive report
H. # external references	Text analysis of descriptive report

In selecting figures of merit, the issue of significance looms… there may be only one interesting image in a collection. Or one key subunit exposure in a kilometer-long traverse. Sometimes highly-valuable fieldwork may be done in a comparatively limited area, while some long-distance traverses return little data of interest. At the risk of introducing subjectivity, as a filter one may consider those features deemed worth describing or reporting by experienced practitioners, as in proposed metrics F-H. These were in some sense "fused" or higher-level measures of scientific merit, compared with simple bit-counts or traverse lengths. In defining the figures of merit, it was therefore decided to drop metrics A-C. Metric E could likewise be just as due to engineering or lighting issues as to the productivity of a given site. The figures of merit then used in this study were metrics D and F-H.

Effects of prior experience, initial conditions

Another issue in conducting a field study is the relative experience of the human participants – both in spacesuits and comprising the rover remote science team – with the geological units surveyed, the processes that formed the region, and the chosen field test sites. Variations in past experience can be expected to affect the interest and questions asked by the scientific participants, and hence affect the scientific return. While this problem remains difficult to eliminate altogether, it can be mollified and evened somewhat through the selection of team members with past experience in the test regions, yet still not allowing the participation of any scientist that has visited that specific test site in the past.

To avoid bias, both the remote robotic and on-site human science teams must be given the same briefing and starting data. In the summer of 2001, ten potential sites were surveyed, collecting remote sensing and aerial imaging for each. A science panel formulated a list of questions for field investigation based on these surveys and reconnaissance, thereby replicating the likely starting point of either a robotic or human EVA investigation [7].

Shirtsleeve geologist as a control

Even with "fresh" participants at a given site, given the same initial data, there remains no absolute yardstick for their respective surveys. Ground truth is typically messy and unfettered. -- we cannot know nearly everything about a given test site without painstaking prior study and analysis. Separate brief surveys of a given test site by robots and EVA could conceivably return two partial sets of different, but equally valid observations that would be difficult to compare. To provide a control to the experiment, a third parallel set of surveys was used – performed by unrestricted human geologists in shirtsleeves, who were given unlimited time to finish.

Need to compare modes of exploration

Humans in EVA and remote-controlled robotic missions operate on different operations timescales. Human endurance and finite life support capacity limits periodic EVAs to a few hours each, while robots may continue on with periodic command downlinks for weeks or months. But robots may effectively waste a large portion of mission/experiment time while awaiting ground instructions (and correcting misinterpreted or off-target instructions), while humans are capable of independent, real-time reprioritization and replanning in the field (e.g., Harrison Schmidt's lunar vitreous "orange soil").

For practical experiment protocol design, some time constraints have to be placed on both human and robotic sorties (except for the shirtsleeved control sorties). In this study, the remote science team operating a given robot could have one or two successive three-hour shifts at a given test site, while humans in simulated EVA were limited to one three-hour shift. Both the rover and suited tests were pre-positioned to a given starting point at each test site. However, like the flexibility given to the "control" geologists, the spacesuited humans were allowed to terminate a field survey prior to the three-hour test limit if they chose, either because their curiosity and objectives had been satisfied or due to suit safety or health concerns.

Procedure

Assume advanced rover capabilities

When considering the future capabilities of robotic missions compared with potential human planetary missions, the relevant timeframe is 10-20 years in the future. Field testing with a current-generation rover such as Nomad [8] or K9 [9] would build-in the limitations and constraints of today's state-of-the-art – not a fair representation of 2015-class robotic exploration systems. Therefore, in this study, it was decided to remotely control a 2015-class rover-equivalent, which was simulated with a modified human-driven all-terrain vehicle, shown in Figure 1 with installed panoramic and targeted imaging capabilities.

Figure 1. Human-operated simulated 2015-class rover at Haughton Crater, Devon Island, Canada.

In a recent study of future space robotic capabilities conducted by Pedersen et al. in 2002 [3], the expected capabilities of 2015-class surface exploration rovers are defined. These include nominal capabilities (no crash programs) of:

- Scientists interact directly, at mission level (not low-level commands)
- Obstacle avoidance and target tracking
- Drive on rough or soft flat surfaces, not steep slopes or boulder fields
- Active rebalancing and/or center-of-gravity control
- Autonomously grasps samples, or can blast/break off a sample
- Self-diagnostics, with preset recovery procedures
- No self-righting or repair
- No auto-map-building or global self-navigation (must be prepositioned)
- Lack human-level cognitive and perceptual capabilities
- Onboard distillation of some science and status data
- Some virtual presence of ground team (visual, not tactile)

The "rover" operator communicated with the remote science team via mission-level text messages and commands to a rover-mounted display. The operator autonomously handled obstacle avoidance, tracking, and rebalancing while driving, and, as directed by the remote science team, could park the rover to acquire samples, use a rock hammer, or take closeup images. But the operator was also instructed to take initiative only in the case of basic safety and recovery procedures (i.e., backing away from a canyon edge or retracing a path out of a loss-of-signal zone), which could reasonably be expected of a 2015-class rover.

Mars-analog field site in Arctic

The site chosen for the July 2002 tests was the ~23Ma Haughton impact structure [10], centered at 75° 23'N, 89° 39'W on Devon Island in the Canadian Arctic. Haughton Crater is a well-preserved structure with an original rim diameter estimated at ~24km [11]. It is an excellent polar-desert Mars analogue that has been shaped by post-impact surface glaciation and periglacial effects, and is nearly devoid of multicelled life. We (the authors) were familiar with this area from past geophysical studies conducted there [12]. An existing NASA-run summer field camp adjacent to the crater provided necessary logistical and scientific support.

Remote science experiment design

Local and trunk wireless networks

As shown in Figure 2, the simulated rover tests required that the vehicle and its operator remain in data communication with a remote science team. Building on previous work in wireless exploration networks [13], field communications with the rover was via a tactical network of 802.11b repeaters, which were in turn connected to base camp by high-speed point-to-point digital spread-spectrum trunk radio links [14]. A commercially-leased 768Kbps satellite link provided connectivity from the field camp to NASA-Ames Research Center, and thence to the remote science team. Including transmission, error-checking and buffering effects, the typical data transmission times from the rover back to NASA-Ames ranged from a few seconds for still images of specimens to 70-90 sec for 3MB 3-color panoramic images.

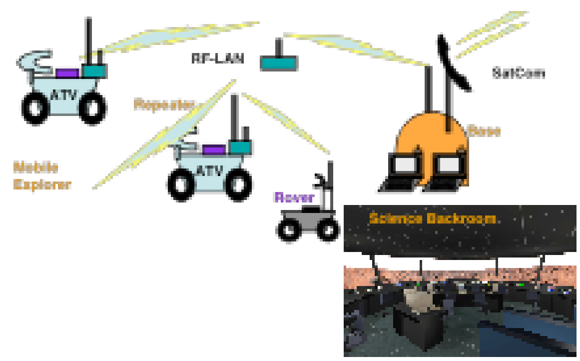

Figure 2. Local and trunk wireless datalinks from the FFC backroom to the rover.

However, the rover tests were run with no added delays, as though they were being operated locally (i.e., from a surface habitat or Mars orbit). This was because it was not feasible to recruit scientists with the patience to voluntarily operate with inserted two-way 20-minute delays to/from the rover. Transmissions to/from the rover were therefore logged with timestamps and a separate post-facto analysis was conducted to construct a similar timeline with the delays inserted.

Virtual presence capability during rover tests

One of the expected capabilities of a 2015-class rover is some degree of virtual visual presence for the remote science team. The Future Flight Central (FFC) facility at NASA-Ames is a full-scale (8m diameter) virtual air traffic control tower with computer-generated projected 360° out-the-window visuals. For this study, the FFC consoles were used to provide image displays to the science team, as well as compose commands for the rover. Panoramic images from the rover were displayed on the FFC "windows", creating a sense of visual immersion for the science team.

Responsibilities for data capture (rover tests, max. 3-hour runs):

- Science team: Personal notes during test run and discussions, combined and turned into a 1-2 page written report afterward. Choose rover-acquired samples up to 5kg limit.
- Support staff: Open communications, archive images by test run, by type, daily
- Rover operator: Trigger scripts to acquire images, dump local image files daily to CDs, clear the onboard storage daily
- Test director: Note stop/start/locations, take GPS waypoints, count total # samples at the end of each run

Tests using HS prototype spacesuit

Prototype spacesuits and geologist test subjects were provided through Hamilton-Sundstrand (HS) and the SETI Institute. The HS suit, shown in Figure 3, was actually an unpressurized "engineering prototype" rather than a flight-worthy suit. Retired Shuttle suit gloves were used with the rigid torso assembly, but there were no leggings. HS provided a crew of two to monitor the safety and health of the geologist test subjects and to transport and maintain the suit.

Figure 3. Hamilton-Sundstrand prototype at Site TA1.

Responsibilities for data capture (human in suit, max. 3-hour runs):

- Geologist in suit: audio recording of personal observations (in suit or external via RF), turned later into a 1-2 page summary; closeup images on camera; choose samples to retain (up to 5kg limit).

- Suit assistant: hand sketches, carry specimen bag, camera and hammer
- Support staff: monitor subject health and safety, download and archive camera images after each deployment, capture handheld video in the field
- Test director: Note stop/start/locations, take GPS waypoints, count total # samples at the end of each run

<u>Shirtsleeved-geologist surveys</u>

Ground-truth surveys by unconstrained geologists were conducted beginning at the same site starting points as in the other test series. An assistant carried cameras, tools and samples.

Written survey reports were received from the science participants after their test runs. Given the figures of merit, tallies of observations, conclusions and hypotheses that were compiled from each report. Figure 4 shows a paragraph from a "raw" report and a corresponding breakout.

Line	Observation	Type	Conclusion	Hypothesis
31			on the inner side of the middle ring/r	
31	side of a hill… covered in brownish scree.	Relational		
32	scree or clasts	Simple		
32	Most ... is covered… clasts	Relational		
32	a broad valley	Simple		
33	standing water	Simple		
32	gently descended ... at the bottom	Relational		
33	polygonal freeze-thaw features	Simple		
34	clasts … in the ...boundaries	Relational		
34	quick-clay type material	Simple		
34	material inside the polygons	Relational		

The general geologic setting is that of a 23 Ma impact into Ordovician-Silurian carbonates. This event gave rise to a concentric series of ridges, valleys, and faults. The test area was on the side of a hill, on the inner side of the middle ring/rim that was covered in brownish scree. Most of the area is covered in scree or clasts. The hill gently descended into a broad valley with standing water at the bottom. There are polygonal freeze-thaw features (larger clasts are concentrated in the polygonal boundaries). There is quick-clay type material inside the polygons. The pebble-to-boulder size scree (or brecciated carbonates) showed signs of weathering. There was amber-streaking at least down to 10 cm depth in some of

Figure 4. Survey report and obsevation tally sheet.

Results

4 remote traverses, 3 suited, 2 free in 7/02

The science team assembled at NASA-Ames at the FFC facility during the week of 22 July 2002. A total of four teleoperated rover test runs were made over a week's time at three separate sites (see Table 2). Weather conditions at the Haughton Crater site were difficult during the week in question: this limited the number of sites visited, time spent at each site, and also slightly affected the quality of the images returned from the four remote runs. Better weather aided in the subsequent visits and traverses to these same three locations by spacesuited (Figure 3) and by unencumbered geologists. However, no descriptive report was turned in by the suited participants at Site T9 or the shirtsleeved visit to TA1.

Table 2. Field tests in July 2002 with teleoperated rover.

ID	Date	Location	UTM-E	UTM-N	Stops	Comments
H024A-1	7/22/02	Site 4	16 417855	8374807	7	Rover commanded into comms hole at 2.5 hrs, "safing procedure"
H029A-1	7/24/02	Site 9	16 421089	8378712	4	Science team wanted multiple pans at each stop
H029A-2	7/25/02	Site 9 (different starting pt)	16 421125	8379050	6	Found water
H0211A-1	7/26/02	Site 11	16 420136	8370769	3*	Surprise starting point "out of landing ellipse" *Lost comms at 1.5 hrs

Discussion

Table 3 lists the science return from each test. As mentioned previously, the times measured for the rover tests did not include Mars-Earth equivalent inserted time delays. Assuming that commands would be grouped whenever possible, post-test analysis indicates that inserting delays into the transmission transcripts increased the typical test duration by a factor of five. If Earth access to a rover were further constrained (i.e., only to typical Deep Space Network twice/day access periods), it is estimated that that would reduce productivity by roughly another factor of five. On the other hand, Mars-controlled rover operations and surface EVAs will be similarly constrained by the daily availability of Mars-based crew.

Table 3. Summary of 2002 field results.

Site	Type	(min) Duration	Observations	Conclusion	Hypothesis
T4	Remote	150	22	4	7
T4	Suited	92	28	4	1
T4	Free	30	41	2	1
T9	Remote	335	18	6	3
T9	Free	7	32	8	3
T11	Remote	90	2	0	0
T11	Suited	63	24	3	2
	Averages-by-type				
	Remote	192	14	3	3
	Suited	78	26	4	2
	Free	19	37	5	2

Of the given metrics, the number-of-hypotheses metric was unexpectedly unreliable – in part because in the relative absence of observations, more speculation was noted (by observing the remote science team) and uncertainties were greater. Total observation count showed the most differentiation between the three test types. The observation averages by test type show, as expected, that human geologists in shirtsleeves are far more productive than either spacesuited humans or teleoperated robots. And spacesuited humans were more productive than the 2015-simulated rovers, with or without adding in the productivity effects of Earth-Mars time delays.

Figure 5 shows the relative observational rates – observations per unit time, normalized.

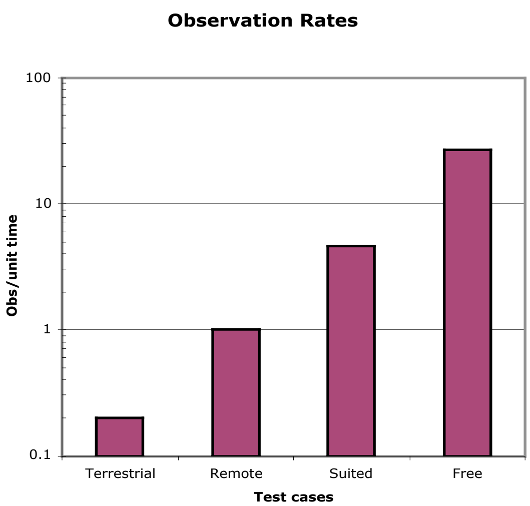

Figure 5. Observations per unit time for the given test cases.

Remote-rovers, terrestrial-controlled (w/delays) observation rate = 0.2
Remote-rovers, Mars-controlled (w/o delays) observation rate = 1 (normalized)
Spacesuited human, observation rate = 5
Shirtsleeve - free human geologist, obs rate = 27

There are too few datapoints to say definitively, but these initial results imply that a local spacesuited human would have nearly 25x the science productivity of a 2015-class Earth-controlled rover. Given that current Mars Exploration Rover (MER)-class systems are only expected to be 5-10% as productive as the 2015-class rovers, this further implies a relative productivity advantage per unit time of around 300-400x for local EVA vs. a current-capability robotic surface mission.

As sketched in Figure 6, it was noted during the teleoperated rover tests that rather than the rover becoming an extension of the remote human science team, the science team became more mechanistic in their planning and execution. Targets of opportunity were bypassed if they were not on the original traverse plan, or would significantly slow the arrival of the rover at its next waypoint. Conversely, both the spacesuited and shirtsleeved humans diverted their traverses to cover nearby targets of interest (such as hydrothermal vents, or a large ejecta block containing macrofossils at site T11).

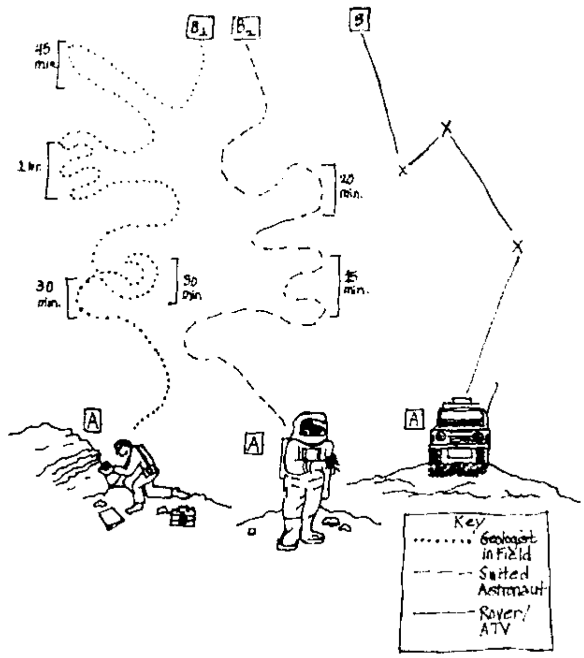

Figure 6. Illustration of rover and humans' paths in field exploration.

Conclusions

Care should be taken to avoid inferring too much from these early results. This study compares human exploration performance to that of a hypothetical 2015 rover at only three test sites in a single geological setting (Haughton Crater). Small changes in mission profile or in the rate of robotic technology maturation can easily skew these figures in either direction. Given that three test points are not enough to be significant, additional field tests of this type are needed.

While human exploration may appear to be 1-2 orders of magnitude more productive than future Earth-controlled robots, the current study seems to indicate that this capability gap narrows with local Mars control of the robots.

Future plans for this study work should address two needs: first, to add more data points to the field science evaluations begun in 2002, going to both a US desert Mars analog site and to either other sites near Haughton Crater or else a second US desert site. This would provide about a total of 8-10 total 3-way science site productivity comparisons between humans and robots, which results would be enough to confidently use in future cost-benefit tradeoffs and other mission planning activities. Second, to examine ways of improving the productivity of teleoperated robots, a separate study could use the same humans in both field exploration and on the remote science team.

Acknowledgements

The authors wish to thank Pascal Lee and the SETI Institute for their field support of this study by the NASA Haughton-Mars Project. Samantha Domville, Victor Rundquist, Richard Alena, and Lori Blaauw all made critical contributions to the success to the 2002 field season and deserve thanks. This study was funded in 2002 by the Human-Robotics Working Group of the NASA Exploration Team (NExT).

References

[1] Friedman, L., "Connecting Robots and Humans in Mars Exploration," *Concepts and Approaches for Mars Exploration*, July 2000, p.118.

[2] Coates, A., "Limited by Cost: The Case Against Humans in the Scientific Exploration of Space," *Earth, Moon, and Planets*, Vol. 87, No. 3, 1999, p. 213.

[3] Pedersen, L., *NASA Exploration Team (NEXT) Space Robotics Technology Assessment Report*, Computational Sciences Division, NASA-Ames Research Center, Moffett Field, California, June 2002.

[4] Spudis, P. and G.J. Taylor, "The Roles of Humans and Robots as Field Geologists on the Moon," *1992 Symposium on Lunar Bases and Space Activities*, NASA Johnson Space Center, 1992, p. 307.

[5] Neal, C., "Geological Investigations of Mars: The Human Factor," *Workshop on Science and the Human Exploration of Mars*, LPI-Contrib-1089, Lunar and Planetary Institute, January 2001, p. 154.

[6]Akin, D., et al., "Robotic Capabilities for Complex Space Operations," AIAA Space 2001 Conference, AIAA-2001-4538, Albuquerque, New Mexico, August 2001.

[7] Snook, K., ed., Remote Science Team Report: Haughton Remote Science Experiment 2002, draft, NASA-Johnson Space Center, November 2002.

[8] Wettergreen, D., et al., "Developing Nomad for Robotic Exploration of the Atacama Desert," *Robotics and Autonomous Systems,* Elsevier, Vol. 26, No. 2, February 1999.

[9] Bresina,J., et al., "Increased Flexibility and Robustness of Mars Rovers," *5th International Symposium on Artificial Intelligence, Robotics and Automation in Space (i-SAIRAS*), Noordwijk, The Netherlands, June 1999.

[10] Jessburger, E.K., "^{40}Ar-^{39}Ar dating of the Haughton impact structure," *Meteoritics*, Vol. 23, pp.233-234, 1988.

[11] Scott, D. and Z. Hajnal, "Seismic signature of the Haughton structure," *Meteoritics*, Vol. 23, pp. 239-247, 1988.

[12] Glass, B. J. and Lee, P., "Airborne Geomagnetic Investigations At The Haughton Impact Structure, " *LPSC XXXII*, Lunar and Planetary Institute, Houston, Texas, 2155.

[13] Gilbaugh, B., B. Glass and R. Alena, "Mobile Network Field Testing at HMP-2000", *2001 IEEE Aerospace Conference*, Big Sky, Montana, March 2001.

[14] Braham, S., et al., "Canada and Analogue Sites for Mars Exploration," *Second Canadian Space Exploration Workshop*, 1999