DOCUMENT RESUME

ED 370 805 SE 054 555

AUTHOR	Hickey, Daniel T.; And Others
TITLE	Using Content-Specific Interest To Evaluate Contemporary Science Learning Environments.
INSTITUTION	Vanderbilt Univ., Nashville. Learning Technology Center.
PUB DATE	Apr 94
NOTE	36p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 1994).
PUB TYPE	Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Astronomy; Classroom Environment; Classroom Research; *Educational Environment; Evaluation Methods; Grade 6; Intermediate Grades; Middle Schools; Science Activities; Science Curriculum; *Science Instruction; Science Programs; Space Sciences
IDENTIFIERS	Attitudes Toward Science

ABSTRACT

This paper describes a framework for studying and evaluating learning environments which contextualize school science content within a larger real-world scientific endeavor, such as carrying on a space mission. A central feature of this framework is its incorporation of recent research on content-specific personal interest. This framework was developed and tested in a pilot evaluation of the Challenger Learning Center's M.A.R.S. (Mission Assignment: Relief and Supply) learning activity. This activity consists of a series of classroom activities which prepare students for a simulated Mars mission at a museum-based learning center. The evaluation involved over 300 students, and provided evidence of the positive impact of this particular program on students' interests, attitudes, knowledge, and activities relative to both science and space science. This evaluation also demonstrated the usefulness of the framework which has been developed for studying contemporary science learning environments. (Author)

Using Content-Specific Interest to Evaluate Contemporary Science Learning Environments

Daniel T. Hickey, Anthony Petrosino, James W. Pellegrino, and

The Cognition & Technology Group at Vanderbilt

Presented at the April, 1994 meeting of the American

Educational Research Association, New Orleans.

The Learning Technology Center

Vanderbilt University

Abstract

This paper describes a framework for studying and evaluating learning environments which contextualize school science content within a larger real-world scientific endeavor, such as carrying out a space mission. A central feature of this framework is its incorporation of recent research on *content-specific personal interest.* This framework was developed and tested in a pilot evaluation of the Challenger Learning Center's *M.A.R.S. (Mission Assignment: Relief and Supply)* learning activity. This activity consists of a series of classroom activities which prepare students for a simulated Mars mission at a museum-based learning center. The evaluation involved over 300 students, and provided evidence of the positive impact of this particular program on students' interests, attitudes, knowledge, and activities relative to both science and space science. This evaluation also demonstrated the usefulness of the framework which we have been developing for studying contemporary science learning environments.

Using Content-Specific Interest to Evaluate Contemporary Science Learning Environments

For several years, researchers have been refining new approaches to instruction which are compatible with the theory of *situated cognition* (Brown, Collins, & Duguid, 1989). This theory holds that the activity, context, and culture in which knowledge is developed is a fundamental part of that knowledge. Application of this theory to the design of learning environments specifies presenting content in the context of "real" tasks with meaningful goals, within a social setting that allows learners to observe others engaged in learning and problem solving (Bransford, Franks, Vye, & Sherwood, 1989; Resnick & Klopfer, 1989). The learning environments which our group is developing follow an approach known as *Anchored Instruction* (Cognition & Technology Group at Vanderbilt, 1990; 1991a). A central feature of anchored instruction is the use of video-based *macrocontexts* to contextualize learning around real-life situations. Ongoing projects such as the *Adventures of Jasper Woodbury* mathematical problem solving series (CTGV, 1990, 1992a, 1992b), the *Scientists in Action* program (Goldman, et. al., 1992; Sherwood, Petrosino, Garrison, Goldman, Hickey, & Bransford, 1993), and the *Mars Mission Challenge* (Hickey, Pellegrino, Petrosino, 1991; Hickey et. al, 1992, & in press) illustrate the applicability of this approach to different instructional domains.

Anchoring school instruction in authentic real world contexts presents new challenges and opportunities for assessment. Controlled studies and classroom observation has shown that anchored instruction programs lead to improved performance on cognitive ability measures including academic skills and problem solving ability (e.g., CTGV, 1992b; Hmelo, et al., 1993; Pellegrino, Hickey, Heath, Rewey, Vye & CTGV, 1992; Van Haneghan, et al., 1992). It is further expected that these environments will leave students with positive affect towards the topics, concepts, and skills which are presented in the environment. Initial studies have shown that anchored instruction environments do, in fact, result in relatively higher levels of interest and self-confidence regarding instructional content (CTGV, 1992b; Hickey, et al., 1993; Sherwood, et. al., 1993).

This paper describes the framework, methods, and results of an evaluation of a learning environment which situates school science content into the larger context of space travel. A more general goal of this research effort was developing and studying a framework for studying and evaluating anchored instruction science programs. While the instruments and techniques developed for this study will be useful for studying other space-related anchored instruction programs (such as our *Mars Mission Challenge*), we expect the instruments to be useful for evaluating most space-related science programs, including the many school science programs sponsored by NASA. Furthermore, we expect the general framework and methods to generalize to any school science program which "situates" academic content within a larger real-world context.

Study Overview

This study evaluated the impact of one of the learning activities offered by the Challenger Learning Center for Space Sciences. This activity, known as *Mission Assignment Relief and Resupply (M.A.R.S),* contextualizes middle-school science content with the larger challenge of a simulated Mars Mission, and includes both classroom and museum-based activities.[1] While not anchored instruction programs *per se,* the Challenger Learning Center's activities include many of the features called for in recent policy recommendations for reforming science curriculum and supported by contemporary educational and cognitive psychology research. This includes incorporating content into a larger meaningful activity stretching over multiple class periods, incorporating realistic hands-on activities, allowing students to select specific topics for more detailed investigation, supporting repeated application of complex concepts in various settings, and creating necessary and realistic needs for communication within and between groups of students.

The *M.A.R.S.* activity consists of a series of classroom "lab" activities and a two-hour simulated mission at the museum-based learning center. At the outset of the "mission", students divide into eight teams (*Communications, Data, Isolation Robotics, Life Support, Medical, Navigation, Probe,* and *Remote*). Each team spends 5 - 10 class periods using carrying out hands-

[1]For more information about the Challenger Center's many space science educational program's, contact them at 1055 N. Fairfax St., Alexandria, VA, 22314, or contact one of the 27 affiliated musuem-based learning centers.

on laboratory activities and learning the concepts and skills which their team will employ during the mission. The museum-based center consists of a "Spacecraft" and a "Base" each equipped with activity stations for each team. The two units are linked by voice, video, and data, and team members must communicate with other members to carry out their teams objectives. As part of the "relief operation," team members switch from the Spacecraft to the Base, and vice versa, halfway through the mission. All eight teams must communicate with each other carry out the mission. (The same hardware and a similar format are used in two other Challenger Center missions, "Return to the Moon", and "Rendezvous with Haley's Comet ".)

The students we observed at the museum-based center were all highly engaged while participating in relatively complex self-directed activities[2]. One might characterize the participants as experiencing the pleasurable psychological state of engagement which Czikzentmihalyi (1990) calls *flow*. This positive emotional experience is expected to leave students more favorably disposed towards the content that they encountered in that environment. Such experiences are expected to positively impact student's interest in the specific science topics which are incorporated, and general interest and attitude towards space travel and space science. To a lesser extent, the experience should also impact interest in and attitude towards specific science topics which were not presented during the mission, and school science in general.

Theoretical Framework

We used the M.A.R.S. evaluation as an opportunity to develop a more general framework for studying how participating in such activities impacts students' *interests, attitudes, activities*, and *knowledge*. Following is an overview of this framework, organized around these four major areas.

Interest. Traditionally, interest has been considered in global terms, typically in terms of interest in "science" in general or in a course topic (e.g., biology, physics), or in science careers. Willson's (1983) meta-analysis revealed a mean correlation of .23 between "interest in science"

[2]We have not, however, observed students engaged in the pre-mission activities in their classroom. We expect a great degree of variability in how materials are actually used by individual teachers.

and science achievement across 33 middle-school studies. This is roughly the same as the mean correlation between socio-economic status and science achievement (mean r =.25, across 21 studies) reported in Fleming and Malone's (1983) meta-analysis of K-12 science studies[3].

In recent years, researchers who study learning and motivation have begun to focus on *content-specific* interest (e.g., Nenninger, 1992; Schiefele, 1991, 1992). An emerging body of research has shown that well-specified conceptions of interest around particular topics are much more predictive of subsequent learning behavior relative to that topic, than more domain-general conceptions of attitude or achievement motivation (e.g., Nenninger, 1992; Schiefele, 1992; Schiefele, Krapp & Winteler, 1992). This emphasis on greater domain-specificity in the study of learning and motivation corresponds with contemporary perspectives in cognition and instruction which underlie anchored instruction. The content-specific perspective on interest is distinguished from other well-known research such as vocational interest (e.g., Campbell, 1974) and curiosity (e.g., Berlyne, 1966) partly by its situation within theories of *action*. Action theories consider interest in terms of how it regulates the individual's choice of actions whenever choices exist (Fink, 1991). This perspective leads to a focus on how interest influences task value, awareness of possibilities for action, and knowledge of possible forms of engagement with objects of interest (Krapp, Hidi, & Renninger, 1992). The importance of these factors in free choice situations makes them highly relevant to studying how individuals learn in environments which depend on self-directed learning.

A predominant theoretical framework in contemporary interest research is the German "Person-Object" theory (Krapp & Fink, 1992, Prenzel, 1992, also Prenzel, Krapp, & Schiefele, 1986, Schiefele, 1986). This theory conceptualizes interest as a unique relation between a person and a class of interest "objects." The theory assumes that individuals assign relatively high value to the goals of interest-oriented action. Krapp and Fink (1992, 406-407) further characterize the relation between a person and interest objects by the presence of *selective persistence* and *self-*

[3]In comparison, Fleming and Malone (1983) report a .43 mean correlation between general ability and science acheivement (42 studies).

intentionality. The relation is associated with positive emotional experience which in turn becomes the center of further action. This theory distinguishes three levels of interest "objects" in terms of specificity (Krapp & Fink, 1992). This includes (1) the most general level of *interest domains,* such as academic subjects and "real world" domains such as space travel, sports, and politics, (2) *interest objects,* specific "things" which are concrete (but not necessarily physical), such as actual problems or challenges, or knowledge which is contained in the learning environment, and (3) *reference objects*, physical objects associated with interest domains and objects. The *Task Interest Survey* which we developed assessed interest in specific learning tasks designed to be representative of specific science topics (i.e., interest domains). We believe that such an instrument can serve as a global measure of science interest (by using score aggregated across domains) as well as a measure of interest in more specific topics. Our task interest instrument also included items in both a general, unspecified real-world context, and in the context of space travel. We expect that students' interest in items presented in the same context as the learning environment would be most useful for studying the impact of participation in that environment

Schiefele (1991, 1992) distinguishes between personal interest as a *latent personality characteristic,* a relatively long-term orientation of an individual towards an interest object, and personal interest as an *actualized state*, of wanting to learn about or become involved with a topic for its own sake ("content-specific intrinsic motivation"). In designing the *Task Interest Survey*, we attempted to operationalize an actualization dimension by assessing interest in general learning activities in a broadly specified domain (low actualization) and in well specified problem solving activities specific to a topic in that domain (high actualization). We expected that more actualized tasks might be most predictive of behavior and more sensitive to an intervention. A central goal of this study was to study the usefulness of each of the three theoretically-based dimensions in creating instruments to evaluate learning environments.

Attitudes towards science. A commonly assessed motivational factor is student's attitude, referring either to affect regarding a domain, or attitude towards learning about that domain in school. Willson's (1983) meta-analysis revealed a mean correlation between attitudes towards science and science achievement (mean $r = .14$, 18 studies) was smaller than the mean correlation between interest in science and achievement (mean $r = .24$, 33 studies) in middle school studies. (This difference, however, was not significant, and the differences between interest and attitude at other school grade levels were smaller.) One objective of this study was to contrast the usefulness of this more conventional type of measure with the task interest instrument which we developed.

Activity. The ultimate goal of activities such as a Challenger Center mission is to lead to desirable behavior in subsequent situations. Ideally, one should be able to link positive changes in interests and attitudes to learning behaviors which, in turn, can be linked to achievement outcomes. We attempted to find out if the *M.A.R.S.* activity leads to increased engagement in a few specific activities, such as discussing or writing about the topics outside of science class, checking out or purchasing relevant books or magazines, watching television programs, etc. In lieu of actually documenting behavior or interviewing students, a simple checklist was used to ask students if they have engaged in specified activities within a certain period of time.

Knowledge. Like anchored instruction environments, the Challenger Center missions are specifically constructed to allow different students to focus on different topics. In the M.A.R.S. activity, students work on one of eight teams, and teachers normally allow students to join the team of their choice. Such individualization presents difficulties for evaluating knowledge acquisition, since students are presumed to acquire different knowledge, depending on team membership. However, we believe that individualization provides an opportunity for evaluating knowledge acquisition not present in more standardized learning environments. In individualized environments, students who focus on a given topic are expected to acquire more knowledge relative to that topic than their classmates who focused on other topics, and vice versa. Thus, students who did not focus on a given topic can be used as a control population for evaluating knowledge gains in the students who did. In the case of the M.A.R.S. activity, we assessed knowledge of concepts presented in the materials used by each of the teams. Posttest scores of students who participated in that team were then compared with those of students on the other teams. Because such a method affords a degree of experimental control for studying knowledge acquisition without requiring the use of untreated control classes, it is potentially useful for evaluating individualized learning environments such as this.

Method

Participants

Students were tested at two sites. At Site 1, one teacher who had participated in previous CLC activities tested all four of her sixth-grade classes who were participating in the M.A.R.S. activity, and recruited four comparison classes taught by four different teachers. At Site 2, one teacher who was leading her first CLC mission tested all four of her sixth-grade classes. While all of her students were participating in a limited set of classroom preflight activities, only about a third of these students were being allowed to participate in actual mission at the CLC. Because only students with "good citizenship" standing were allowed to participate in the mission simulation at the CLC, these two groups are not equivalent. There were no control students at Site 2. Thus, the total sample (308 students) consisted of eight experimental classes participating in the *M.A.R.S.* activity, and four control classes[4].

Procedure

CLC teachers were instructed to test students at least one week before beginning any mission-related activities[5] and three weeks after attending the Challenger Center.[6] Control teachers were instructed to administer the instruments on the same days as the CLC teachers, and were asked to defer from any space-related instruction during the six-week period. A brief set of instructions were provided for the teachers to read to the students. These instructions explained the

---

[4]The instruments were also administered to 500 seventh-graders in Nashville, TN, to provide a sample for further examining the theoretical dimensions underlying the instruments. These findings are reported in Hickey, Petrosino, & Pellegrino (1994).
[5]CLC students at both sites were familiar with the Challenger Center and were already aware that they would be participating in a mission at pretest.
[6]This three-week delay may have attenuated posttest scores in the CLC classes. However such a delay was necessary to minimize the tendency for students to include activities carried out as part of the intervention in their responses to the activity survey(which asked about activities carried out in the previous two weeks).

study to the students and encouraged them to be "honest" and to "do their best". Additional instructions were provided on each instrument. The pretest and posttest administrations each occupied an entire class period.

Analyses

In the course of the study, we found that nearly half of the Site 1 CLC students and nearly one quarter of the Site 1 control students reported participating in CLC missions during previous school years. Given the obvious confound of previous CLC activity, Site 1 results were analyzed two ways. First, the pretest scores of all students were used to contrast students who reported participating in CLC activities with the students who did not[7]. The second contrast examined Site 1 pretest to posttest change only in the students who did not participate in previous missions. Site 2 results were analyzed by contrasting the pretest to posttest change in the students who actually attended the museum-based center with their classmates who did not.

Instruments.

Seven instruments were developed for this study. Three of them (the interest, attitude, and activity surveys) would be appropriate for evaluating any learning environment that presents school science in the context of space travel, while the other four were specific to the M.A.R.S. activity These instruments were administered using machine-scoreable forms.

Task Interest Survey. This instrument assessed students' interest in 40 well-specified school science tasks. Students were asked "How do you think you would feel" regarding each task, responding on a six-point scale ranging from *very bored* to *very interested.* Following the logic described above, these tasks were developed along the dimensions of *content, context, and actualization.*

The *content* dimension represents different school science domains. To identify an objective set of topics, we reviewed the middle school objectives from the *1990 Tennessee State*

[7]Because the Site 1 students attended different elementary schools during the previous year, we were unable to determine how students came to participate in prior CLC activities. The role of self-selection by *individuals* is small, since participation is with one's school class. Self-selection by the *teachers* does present a possible confound, because participation in some some schools is left up to the teacher. Thus students who were more interested in science and space or more academically oriented may have been more likely to attend the Challenger Center.

Science Framework. Eight specific topics from three areas (Physical, Life, and Earth Sciences) were selected (Table 1). The *context* dimension represented the distinction between tasks presented in a general context and a space travel context (see Table 2). The *general* context items presented tasks contextualized in a general "real world" context while the *space* context tasks were presented in the context of space travel, such as students might encounter in a space-oriented anchored instruction environment.

The third interest dimension, *actualization,* was operationalized with tasks that involved either *Learning about* a topic area in general, *Reading about* a more specified topic in that area, or *Figuring out* a specific problem derived from that topic. Table 2 lists the five items from the *Life Science* topic *Plants.* We choose not to include *Learning about.... space* context items because this combination tended to yield nonsensical tasks such as *Learning about plants in space.* (In other words, based on our definitions, it appears that contextualizing tasks also served to increase their actualization.) As shown in Figure 1, this configuration yielded 16 space context items and 24 general context items. The entire set of items is included in Appendix 1.

Our purpose in constructing such an interest instrument was to explore whether the intervention differentially impacted specific topic areas, and whether different types of tasks were more sensitive to instructional interventions. In particular, we expected the more highly actualized items (*Figuring out...*) and the space context items to be more sensitive to the instructional intervention.

Attitude Survey. We used a seven-item "Attitude Towards Science" scale developed by Ebenezer and Zoller (1993), shown in Table 3. Pilot test results revealed that two of the items from this scale (*I feel that it is important to study science*, and *Science is a valuable subject*) were poorly correlated with the other five items (e.g., *I like to study science in school, Science classes are boring*). It appeared that these two items assessed "value" attached to learning science, while the other five appeared to capture students "feelings" about learning science. Thus the two value items, plus three new ones, were used to create a separate scale, yielding a five item *value* subscale and a five item *feeling* subscale. As shown in Table 3, each of these items were rewritten to assess students attitudes towards the domain of space and space travel, yielding the four subscales which make up the *Space/Science Attitude Survey.*

Activity Survey. Based on a format used by Skinner and Barcikowski (1973), the *Space/Science Activity Survey* asked students whether or not they engaged in various science-related activities over a specified period of time. As shown in Table 4, these activities are grouped into three areas: 1) general science, 2) space travel and space science, and 3) science topics which are explicitly presented in the *M.A.R.S.* activity. Students are asked how many times (0 to 5+) they engaged in each activity *outside of science class* within the previous two weeks.

Knowledge Survey. In order to assess knowledge acquisition, a 20 item multiple-choice test was constructed. For six of the eight teams, three multiple choice items were constructed based on the materials used by each team in the classroom "preflight" activities.[8] Two additional items assessed science concepts which were presented in lab activities completed by all students regardless of team. The items were designed to be quite difficult to prevent a ceiling effect. Each item had six multiple-choice response, one of which was *not sure,* and the instrument instructed students to not guess at the correct answer. The subject and format of each item is presented with the results in Table 5.

M.A.R.S. Activity Survey. This instrument was used to study Interestingess of the 20 classroom "preflight" learning activities. CLC students were asked to check whether or not they participated in each activity, and to report whether the activity made them feel "bored" or "interested" on a six-point scale. Students were also asked to rate the interestingness of the preflight activities in general, and the "mission control" and "spacecraft" components of the museum based activity.

Background Questionnaire. This instrument asked students whether they participated in previous Challenger Center missions or NASA's "Space Camp", and if their parents or adult acquaintances work for NASA, in the aerospace industry, or as scientists. Students were also

[8]No items specific to the *Data* and *Communication* teams were constructed because of space restrictions and the lack of clearly defined science topics in the materials for these two teams

asked how often they watch any of the Star Trek television shows, and what other space-related activities they participated in. A posttest questionnaire asked CLC students which team they were on and whether they actually participated in the preflight and CLC activities.

Results

General and Background

Examination of the student responses revealed that most students were in fact able to complete the instruments in a single class period. Only 6 of the respondents failed to complete the forms. Table 4 shows the number of participants in each of the three groups at pretest and presents the information obtained from the *Background Questionnaire*. In terms of gender, exposure to adult scientists, and exposure to *Star Trek*, the groups appear quite similar.

Interest

Each of the eight five-item content-specific interest scales was highly reliable, with internal consistencies (Cronbach's Alpha) of .80 or greater. Figure 2 compares the pretest scale means for the Site 1 students. While the scores are higher for the students who participated in previous Challenger Center activities, the difference approached significance only in the Astronomy and Geology scales, $F(1,194) = 2.5$, $p = .11$, and $F(1,194) = 2.6$, $p = .11$, respectively.

Pretest and posttest scores were examined for group-by-time interactions, which indicate differences in *change* between the CLC and control students from pretest to posttest. At Site 1, eliminating students who participated in CLC activities during the previous school year left only 26 experimental students and 79 control students. Comparing these two groups revealed significant effects of group (experimental students were higher on most of the eight scales at both pretest and posttest) but no significant group-by-time interactions. At Site 2, complete pretest and posttest data was available for only 24 students who participated in the classroom preflight activities *and* attended the Challenger Center, and 44 students who participated in the classroom activity but *did not* attend the Challenger Center. While attendees at Site 2 were higher on all eight scales, there were no significant group-be-time interactions.

In order to determine if different types of items were more sensitive to group differences, the task interest items were collapsed along the context and actualization dimensions, yielding five eight-item scales. As shown in Figure 3, the difference for the Site 1 Previous-CLC and No-Previous CLC groups does in fact appear largest for Figuring Out/Space Context items, (but none of these differences were statistically significant, Figuring Out/Space, $F\ (1,194) = 1.7$, ns; Reading/Space, $F\ (1,194) = 1.1$, ns; others, $F < 1$). Thus it appears that the different types of interest tasks do in fact differ in their sensitivity to meaningful group differences, and that highly actualized tasks which matched the context of the intervention appear to be most sensitive.[9]

Attitudes.

All four attitude scales were highly reliable. Internal consistencies (Cronbach's Alpha) on the Science Feeling, Science Value, Space Feeling, and Space Value scales were .86, .87, .84, and .84, respectively. Figure 4 contrasts the scale means for the Site 1 students with previous CLC experience to the Site 1 students with no CLC experience. Students with previous CLC experience were higher on all four scales, but the difference was statistically significant difference only on the two space scales; Science Feeling, $F(1,194) = 2.9$, ns; Science Value, $F(1,194) = 2.2$, ns; Space Feeling, $F(1,194) = 12.0$, $p < .001$; Space Value, $F(1,194) = 7.9$, $p < .005$.

Examination of the pretest and posttest attitude results at Site 1 revealed that the experimental students were significantly higher on both of the Space Attitude scales at both times, but there were no significant group-by-time interactions. Figures 4a thru 4d contrast Site 2 pretest-posttest changes between the student who only participated in the classroom preflight activities and the students who participated in the preflight activities *and* attended the CLC. While the two Value scales show more positive change in the students who attended CLC, the interactions were not significant: Science Value, $F(1,66) = 1.4$; Space Value, $F(1,66) = 2.0$.

[9]The correlations among the discrete groups of items here were high (over .80) suggesting limited discriminabilty for the context and actualization dimensions. This issue is explored in more detail using a larger sample in Hickey, Petrosino, & Pellegrino (1994).

Activity

Scores on the Space/Science Activity Survey were analyzed by collapsing the results for each activity into two categories (0 times and 1 or more times). Figure 5a contrasts the reported pretest science-related activities for previous-CLC/no previous-CLC groups at Site 1. The Chi-square test of these differences was statistically significantly ($p < .05$) for items 2, 3, and 4. Figure 5b contrasts the space-related activities for this same population. The differences for were statistically significant ($p < .005$) for all except item 12. Figure 5c contrasts the differences for this population on the *M.A.R.S.*-related items, with significant difference for items 13, 16, 18, and 20 ($p < .005$).

Examination of the Site 1 pretest/posttest activity results revealed differences between the experimental and control students on two of the Space activity items. As shown in Figure 7, more of the experiment students reported writing and drawing about space outside of science class at posttest, compared to the control students. The differences between the groups were much smaller at pretest. Unfortunately, non-normally distributed frequency data from repeated measures are difficult to analyze statistically. Furthermore, the Site 1 results are confounded by the fact that both experimental and control students were pretested the day after returning from Spring vacation.

Knowledge

All students at Site 1 completed the *Knowledge Survey* instrument at posttest. The items were shown to be quite difficult. with the proportion of students selecting *not sure* exceeding the proportion of correct answers on many items. Table 5 list the content and format of each item and the shows the percentage of experimental and control students who reported the correct answer for each item. While the experimental students outperformed the control students on most items, the difference was statistically significant only for items 2, 6, 9, 12, 14, and 15 (Chi-square test, all $p < .02$). Thus there is evidence for meaningful advantage in content knowledge in students who participated in the M.A.R.S. activity.

Since the knowledge items were derived from materials used primarily by members of one team, we expected members of specific teams to register a higher score for the items which were specific to their team, compared to the items which were specific to the other teams. Aggregate scores for the six sets of three team-specific items were computed. Table 6 presents the mean number of items correct for each team on the set of team-specific items. No clear pattern of advantage for the team matching the item sets emerged, suggesting that being a member of the teams did not lead to greater learning of the team-specific content. However, the very small sample size and the difficulty of the items seriously qualifies the strength of these conclusions. This example does, however, illustrate a potentially useful manner for studying knowledge gains in individualized learning activities such as this one.

Interestingness of M.A.R.S Activities

The *M.A.R.S. Activity Survey* was administered only to the experimental students at Site 1. The Site 1 M.A.R.S. teacher confirmed that the students in her class completed all of the lab activities in the teacher's manual and felt that her students would recall participating in most of the activities when reading the labels used on the survey. Figure 8 presents the number of students who rated each activity and the mean rating for each. It appears the students found the *Lost on the Moon, Mission Emblems, Communication Lab,* and *Robots & Job-Bots* activities relatively more interesting, while they found the *Graduated Cylinder, Martian Canals, Holiday in Space* and *Hello Out There* activities relatively less interesting.

We also considered the interestingness of the different activities by team membership. The Site 1 students who participated in the M.A.R.S. activity completed a *Posttest Questionnaire* which had them indicate which team they were on and asked them to rate the interestingness of each activity on a six point scale (from *very boring* to *very interesting*). As shown in Figure 9, students reported very high levels of interest in all three components. The differences between Preflight and Spacecraft ratings and between Spacecraft and Base ratings were statistically significant ($p < .005$). Thus, while the students rated all three activities quite highly, they found the Spacecraft activity more interesting than either Preflight or Base activities.

Since different teams engaged in different tasks during each of the three major activities, these results were broken down by team membership. Figures 10 display the mean interestingness for the three parts of the M.A.R.S. environment for each of the eight teams. Substantial differences are present for the different teams on both the Preflight and Base activities, but not the Spacecraft activity[10]. Given the small sample for the individual team ratings, these finding should be treated cautiously. With less that 10 individuals reporting for each team, one individual can strongly influence each team's mean rating. We consider this to be an example of another potentially useful way of studying highly individualized learning activities such as this one.

Discussion

Summary of the Evaluation Framework

As a pilot study of an evaluation framework for an innovative learning environment, this study was quite successful. In particular, the attitude and interest instruments demonstrated high internal consistency, and teachers at remote sites were able to administer the entire set of instruments within a single class period. Only a few students failed to complete the instruments or provided questionable responses. We feel that all of the instruments developed for this evaluation are potentially useful instruments for evaluating and studying contemporary science learning environments.

The task interest instrument did in fact appear sensitive to substantively different types of student interests. As we had expected the most highly actualized space context items appeared most sensitive to the group differences. However at the global level, this instrument was not as sensitive to group differences as the more conventional attitude measure.

This pilot study also pointed to key issues for any large-scale evaluation of the M.A.R.S. activity. The lack of equivalence in CLC and control students at pretest confirms the need for a pretest-posttest design. This study also demonstrates a likely confound presented by students who participated in CLC in previous school years, and illustrates the need to carefully schedule pretesting and posttesting times. Due to the high number of Site 1 students who reported previous

[10]Note that scores are near ceiling for all teams on the Spacecraft activity. There may have been differences between teams that were beyond the range of this scale.

CLC participation, the number of experimental students was quite small. Additional evaluation studies with larger samples and more control over administration will help us understand the usefulness of these instruments in pretest-posttest studies.

Impact of the M.A.R.S. Activity.

As an initial evaluation of the impact of the M.A.R.S. activity, these results are positive. In particular, we found that:

1. Compared to their classmates, students at Site 1 who reported participating in CLC activities in previous school years reported (1) higher interest in and value for learning about science and space travel, (2) higher personal interest in seven of eight middle school science topics, and (3) ,ore engagement in science and space-related activities outside of science class.
2. Compared to the control students, the Site 1 students who participated in the CLC M.A.R.S. activity reported (1) more writing and drawing about space outside of science class, and (2) greater knowledge of topics presented in the M.A.R.S. activity.
3. Compared to their classmates who only participated in the classroom preflight activities, the Site 2 students who attended the Challenger Center reported more positive change in their beliefs regarding the value of learning about science and space travel.

Evaluation of Specific M.A.R.S. Activities

Parts of this study examined the interestingness of different components of the M.A.R.S activity. The CLC students at Site 1 reported very high levels of interest in all aspects of the activity. Regarding the various specific activities, CLC students reported:

1. The *Communications* and the *Job-Bots* preflight classroom activities were relatively more interesting, while the *Graduated Cylinders* and *Holiday in Space* activities were relatively less interesting.
2. The CLC *Spacecraft* activity was more interesting than the CLC *Mission Control* activity, which was more interesting than the classroom preflight activities.

3. There appear to be meaningful differences in the interestingness of the preflight and Mission Control activities, depending on which of the eight teams the students belonged to.

Acknowledgments

This activity was supported in part by the Tennessee Space Grant Consortium, which is funded by the NASA Space Grant College & Fellowship Program. Thanks to Dr. Alvin Strauss, TSGC Director, and Gayle Ray, Coordinator. Thanks also to Amy Bordeaux, Vice President for Education, Challenger Center for Space Sciences, Gene Nibbelink, Tampa Challenger Center Flight Director, Betty Glass, Houston Challenger Center Flight Director, and Gene Snow, Bridgeport Challenger Center Flight Director, for feedback on study design and coordination of survey administration. Special thanks to the five teachers whose classrooms participated in this study (especially Deborah Allen of Burney Simmons School, Plant City, FL) and their students.

Other members of the Vanderbilt Learning Technology Center who contributed to this project include John Bransford, Susan Goldman, Dan Schwartz and Bob Sherwood. The opinions expressed here do not necessarily reflect those of the Tennessee Space Grant Consortium, NASA Space Grant College & Fellowship Program, Vanderbilt University, or the Vanderbilt Learning Technology Center. Correspondence concerning this study should be addressed to Dan Hickey, Box 45, Peabody College, Nashville, TN 37203. Electronic mail may be sent to hickeydt@ctrvax.vanderbilt.edu.

References

Berlyne, D. E. (1966). Curiosity and exploration. *Science,153*, 25-33.

Bransford, J. D., Franks, J. J., Vye, N. J., & Sherwood, R. D. (1989). New approaches to instruction: Because wisdom can't be told. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 470-497). New York: Cambridge University Press.

Campbell, D. P. (1974). *The Strong-Campbell Interest Inventory.* Stanford CA: Stanford University Press.

Cognition and Technology Group at Vanderbilt. (1990). Anchored instruction and its relationship to situated cognition. *Educational Researcher, 19(3),* 2-10.

Cognition and Technology Group at Vanderbilt. (1991a). Technology and the design of generative learning environments, *Educational Technology, 31(5)*, 34-40.

Cognition and Technology Group at Vanderbilt (1991b). Some thought about constructivism and instructional design. *Educational Technology, 31, (9)*, 16-19.

Cognition and Technology Group at Vanderbilt (1992a). The Jasper experiment: An exploration of issues in learning and instructional design. In M. Hannafin & S. Hooper (Eds.), *Educational Technology Research and Development.*

Cognition and Technology Group at Vanderbilt (1992b). The Jasper series as an example of anchored instruction: Theory, program description, and assessment data. *Educational Psychologist.*

Csikszentmihalyi, M. (1990). *Flow--the psychology of optimal experience.* New York: Harper & Row.

Ebenezer, J. V., & Zoller, U. (1993). Grade 10 students' perceptions of and attitudes towards science teaching and school science. *Journal of Research in Science Teaching, 30 (2)*, 175-186.

Fink, B. (1991). Interest development as structural change in person-object relationships. In L. Oppenheimer and J. Valisner (Eds.), *The origins of action: Interdisciplinary and international perspectives.* New York: Springer-Verlag.

Fleming, M. L., & Malone, M. R. (1983). The relationship of student characteristics and student performance as viewed by meta-analytic research. *Journal of Research in Science Teaching, 20 (5),* 481-495.

Goldman, S. R., Petrosino, A., Sherwood, R. D., Garrison, S., Hickey, D., Bransford, J. D., & Pellegrino, J. (in press). Multimedia environments for enhancing science instruction. in S. Vosniadou, E. De Corte, R. Glaser, & H. Mandl (Eds.), *International perspectives on the psychological foundations of technology-based learning envrionments.* Springer-Verlag

Hickey, D. T., Pellegrino, J. W., & Petrosino, A. (1991, October). *Reconceptualizing space science education: A generative, problem-solving approach.* Paper presented at the Florida Space Education Conference, Cocoa Beach, FL.

Hickey, D. T., Pellegrino, J. W., Goldman, S. R., Vye, N. J., & Moore, A.L. (1993, April). *Interests, attitudes, & anchored instruction: The impact of one interactive learning environment.* Paper presented at the Annual Convention of the American Educational Research Association, Atlanta, GA.

Hickey, D. T., Petrosino, A. Pellegrino, J. W, & CTGV (1993). *Challenger Learning Center M.A.R.S. Activity Pilot Evaluation Study.* Unpublished Manuscript. Vanderbilt Learning Technology Center.

Hickey, D. T., Petrosino, A. Pellegrino, J. W, & CTGV (1993, April). *Middle schoolers' interest in science and space: Dimensions of content, context, and actualization.* Paper presented at the meeting of the American Educational Research Association, New Orleans.

Hickey, D. T., Petrosino, A. Pellegrino, J. W. Bransford, J. B.. Goldman, S. R., & Sherwood, R. (in press). The Mars mission challenge: A generative problem-solving school science environment. In S Vosniadou & H. Mandl (Eds.) *Psychological and educational foundations of technology-based learning environments* (NATO ASI Series). Berlin: Springer-Verlag

Hickey, D. T., Petrosino, A. Pellegrino, J. W. Bransford, J. B., Goldman, S. R., & Sherwood, R. (1992, April). *The Mars mission challenge environment: A generative problem-solving school science environment.* Paper presented at the NATO Advanced Studies Institute Workshop on the Psychological and Educational Foundations of Technology-Based Learning Environment. Koiymbari, Crete.

Krapp, A. & Fink, B (1992). The developmental function of interests during the critical transition from home to preschool. In K. A. Renninger, S. Hidi & A. Krapp, A. (Eds.). *The role of interest in learning and development.* Hillsdale: Earlbaum, 71-98.

Krapp, A., Hidi, S., & Renninger, K. A. (1992). Interest, learning, and development. In K. A. Renninger, S. Hidi & A. Krapp, (Eds.). *The role of interest in learning and development.* Hillsdale: Earlbaum, 71-98.

Nenninger, P. (1992). Task motivation: An interaction between the cognitive and content-oriented learning dimensions. In K. A. Renninger, S. Hidi & A. Krapp, A. (Eds.). *The role of interest in learning and development.* Hillsdale: Earlbaum, 121-150.

Pellegrino, J. W, Hickey, D T, Heath, A., Rewey, K., Vye, N., & CTGV (1992, April). *Assessing the outcomes of an innovative instructional program: The 1990-1991 implementation of "The Adventures of Jasper Woodbury".* Presentation. Annual Meeting of the American Educational Research Association, San Francisco, CA.

Prenzel, M. (1992). The selective persistence of interest. In K. A. Renninger, S. Hidi & A. Krapp, A. (Eds.). *The role of interest in learning and development.* Hillsdale: Earlbaum, 71-98.

Prenzel, M., Krapp, A. & Schiefele, U. (1986). Grundzüge einer pädagogischen Interessentheorie. *Zeitschrift für Pädagogik, 32,* 163-173

Renninger, K. A. (1992). Individual interest and development: Implications for theory and practice. In K. A. Renninger, S. Hidi & A. Krapp, A. (Eds.). *The role of interest in learning and development.* Hillsdale: Earlbaum, 361-397.

Resnick, L. G. & Klopfer, L. E. (Eds.) (1989). *Toward the thinking curriculum: Current cognitive research.* Alexandria, VA: American Society for Curriculum and Development.

Schiefele, H. (1986). Interesse-Neue Antworen auf ein altes Problem. *Zeitschrift für Pädagogik, 32,* 153-162.

Schiefele, U. (1991) Interest, learning, and motivation. *Educational Psychologist, 26,* 299-323.

Schiefele, U. (1992). Topic interest and levels of text comprehension. In K. A. Renninger, S. Hidi & A. Krapp, A. (Eds.). *The role of interest in learning and development.* Hillsdale: Earlbaum, 151-182.

Schiefele, U., Krapp, A. & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of the research. In K. A. Renninger, S. Hidi & A. Krapp (Eds.), *The role of interest in learning and development.* Hillsdale: Earlbaum, 183-213.

Skinner, R., & Barcikowski, R.S. (1973). Measuring specific interests in biological, physical, and earth sciences in the intermediate grade levels. *Journal of Research in Science Teaching, 10 (2),* 153-158

Sherwood, R., Petrosino, A., Goldman, S. R., Garrison, S., Hickey, D. T., Bransford, J.D., Pellegrino, J.W. (1993, April). *An experimental study of a multimedia instructional environment in a science classroom.* Presentation at the meeting of the American Educational Research Association, Atlanta, GA.

Van Haneghan, J., Barron, L., Young, M., Williams, S., Vye, N., & Bransford, J. (1991). The Jasper Series: An experiment with new ways to enhance mathematical thinking. In D. Halpern (Ed.), *Concerning the development of thinking skills in the sciences and mathematics.* American Association for the Advancement of Science.

Willson, V. L. (1983). A meta-analysis of the relationship between science achievement and science attitude: Kindergarten through college. *Journal of Research in Science Teaching, 20 (9),* 839-950.

Table 1
Middle School Science Topics [a]

Science Area[b]	Topic
Physical Science[c]	Machines & Work Electricity & Magnetism Sound, Heat, & Light
Life Science[d]	Animals Plants The Human Body
Earth/Space Science[e]	Astronomy Geology

[a]Based on grades 6-8 objectives for the *Tennessee State Science Framework.*
[b]The area *Environmental Science* was excluded.
[c]The area *Matter and Energy* was excluded.
[d]The topics *Growth and Development* and *Microscopic Life* were excluded from this area.
[e]The topics *Meteorology* and *Oceanography* were excluded from this area.

Table 2
Example Task Interest Items (Life Science/Plants)

Task	Context	Item[1,2]
Learning	General	Learning about plants.
Reading	General	Reading about how plants grow in greenhouses
	Space	Reading about how plants grow in the weightlessness of space.
Problem Solving	General	Figuring out how many trees are needed to make enough oxygen for one person.
	Space	Figuring out how many trees are needed to make enough oxygen for a Mars colony.

[1]All items prefaced with the statement *How do you think you would feel...*
[2]Scored on a 1-6 scale of agreement: *Very Bored, Bored, Slightly Bored, Slightly Interested, Interested, Very Interested.*

Table 3
Science/Space Science Attitude Survey Items

Topic	Valence	Items
Science	Feeling	I like to study science in school.[a]
		Science is dull.[a]
		I do not enjoy science.[a]
		I would like to study more science.[a]
		Science classes are boring.[a]
	Value	I feel that it is important to study science.[a]
		Science is a valuable subject.[a]
		Learning science is not very important to me.
		I don't feel it is very important to study science.
		I don't think science is a very valuable subject.
Space	Feeling	I like to learn about space and space travel.
		Space science is a dull topic.
		I do not enjoy learning about space travel.
		I would like to learn more about space.
		Learning about space is boring.
	Value	I feel that learning about space is very important.
		Space and space travel are not very important to me.
		Space science is a very valuable subject.
		I don't think it is important to learn about space travel.
		I don't think space travel is a very valuable topic.

Note: Items scored on a 1-6 scale of agreement: *Strongly Disagree, Disagree, Barely Disagree, Barely Agree, Agree, Strongly Agree.*
[a]From Ebeneezer and Zoller (1993).

Table 4
Science/Space Science Activities Checklist Items

Category	Items
General	Check out a library book (each book counts as one time)?
Science	Check out a library book about science? Buy a magazine or book about science? Write anything about science outside of science class? Watch a television program about science at home? Talk about science outside of science class?
Space Science	Talk about or read about space travel at home? Check out a library book about space or space travel? Buy a book or magazine about space or space travel? Draw anything about space or space travel at home? Write anything about space or space travel outside of science class?
Specific M.A.R.S. topics	Talk about or read about communication? Talk about or read about radioactivity? Talk about or read about how plants grow? Talk about or read about how computers work? Talk about or read about how the human body works? Talk about or read about the constellations of stars? Talk about or read about electricity and electrical circuits? Talk about or read about robots or robotics?
Foil	Talk about history outside of history class.

Note: All items are prefaced with the stem: *In the last two weeks, did you....* Instructions tell student to only count activities done outside of science class. Response categories are *0 times, 1 time, 2 times, 3 times, 4 times, 5+ times.*

Table 5
Participants (at Pretest)

	Site 1 CLC	Site 1 Control	Site 2
Number	83	116	89
Any Previous CLC Mission?			
Yes	38	27	6
No	45	89	83
Gender			
Male	57%	50%	46%
Female	43%	50%	54%
Demographics			
Attended Space Camp	5%	0%	3%
Know adults who work for NASA	8%	14%	11%
Parents work for NASA	0%	0%	0%
Know adults in space industry	8%	10%	7%
Parents work in space industry	2%	1%	1%
Know a scientist	12%	17%	7%
Parents are scientists	4%	2%	0%
Watch any *Star Trek* shows:			
"Never"	38%	44%	40%
"Sometimes"	32%	37%	38%
"Often"	12%	8%	10%
"Every Chance I Get"	18%	11%	11%
Other Space-Related Activities[11]			
None Listed			63%
One Listed			33%
Two Listed			4%

[11] Site 1 students were not asked this question.

Table 6
Percentage Knowledge Items Correct at for CLC and Control Students.

Item[a]	CLC[b]	Control[c]	Item	CLC	Control
1. Density (source) (ISO)	25.8	14.8*	11. Transceiver (function) (PR)	10.1	13.3
2. Barometer (function) (LS)	32.6	18.0*	12. Color of Mars (iron) (REM)	41.6	9.4***
3. Respiration (purpose) (MED)	49.4	32.8*	13. Evaporation (result) (ISO)	12.4	15.6
4. Azimuth (label) (NAV)	5.6	7.0	14. Hygrometer (function) (LS)	12.4	7.8
5. Gyroscope (function) (PR)	7.9	7.8	15. Weightlessness (result) (MED)	10.1	4.7
6. Photosynthesis (product) (REM)	9.0	7.0	16. 20°E, 20°S (locate) (NAV)	6.7	9.4
7. Radium Decay (product) (ISO)	5.6	5.5	17. CPU (function) (PR)	14.6	4.7**
8. Base (function) (LS)	27.0	7.0***	18. Balance (function) (REM)	18.6	16.4
9. Receptors (example) (MED)	9.0	6.3	19. Millimeters (in meter)	25.8	18.0
10. Lat & Longitude (find) (NAV)	3.4	1.6	20. Gravity (function)	16.9	16.4

[a]Type of item and CLC team are presented in parentheses. Each item had five multiple choice responses, plus *not sure.*
[b]N = 89
[c]N = 128
χ^2, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 7.
Mean correct score of team-specific knowledge items for team members at Site 1.

Team	Items	Team (No. Members)					
		ISO (9)	LS (11)	MED (12)	NAV (12)	PR(10)	REM(11)
Isolation-Robotics (ISO)	1, 7, 13	**.55**	.27	.58	.33	1.0	.27
Life-Support (LS)	2, 8, 14	.55	**.27**	.83	.66	.60	.36
Medical (MED)	3, 9, 15	.33	.72	**.58**	.58	.70	.45
Navigation (NAV)	4, 10, 16	.11	.27	.00	**.16**	.30	.00
Probe (PR)	5, 11, 17	.22	.27	.33	.08	**.60**	.18
Remote (REM)	6, 12, 18	.77	.27	.58	.66	.80	**.36**

Note. Maximum score = 3. Scores on diagonal (bold) represent match between team-specific item means and scores for member of that team.

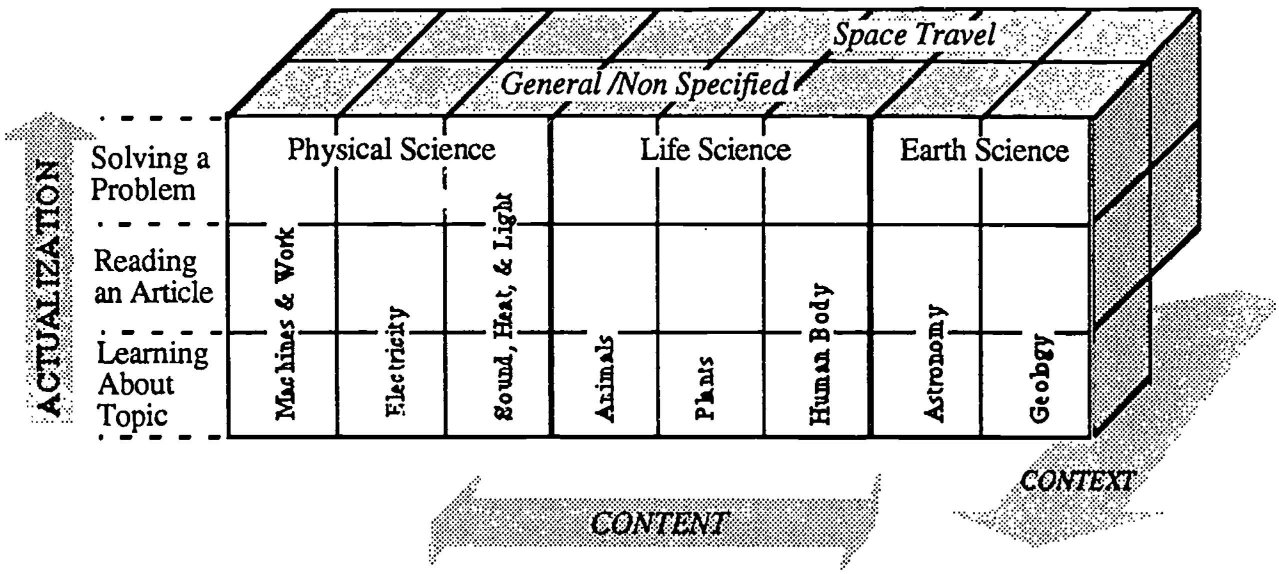


Figure 1
Three dimensions of content-specific interest measured.

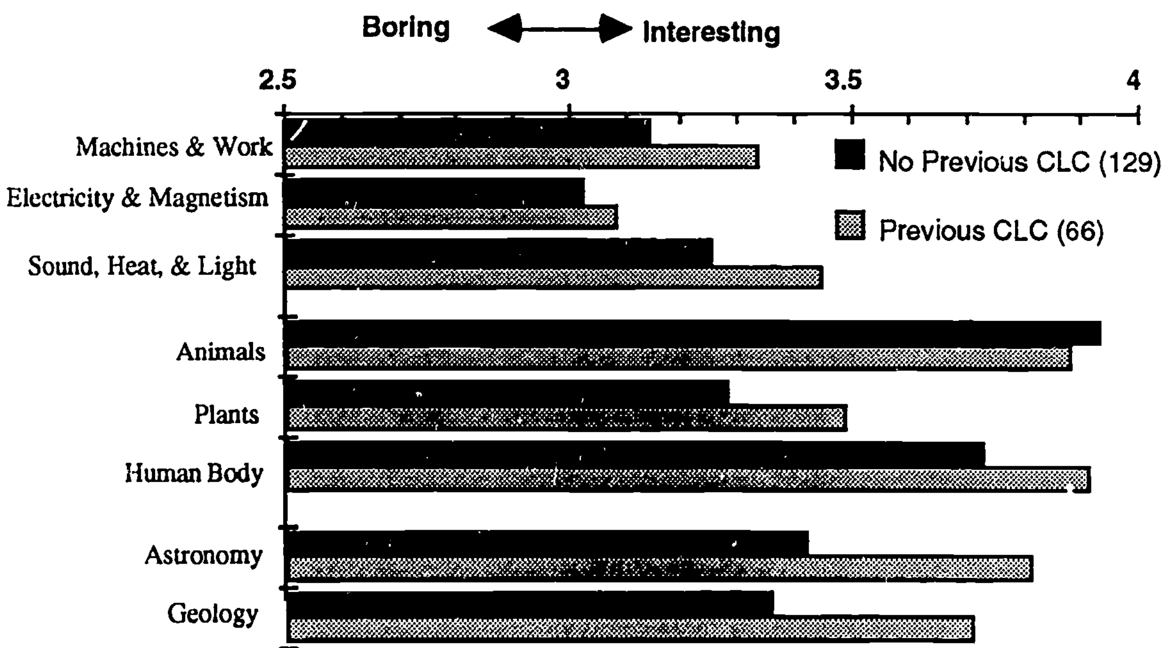


Figure 2
Pretest topic-specific interest for students who did and did not participate in CLC during previous school year at Site 1.

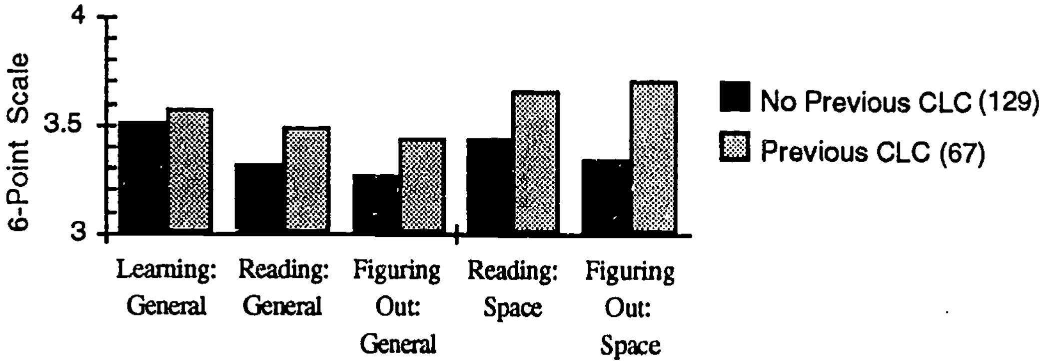


Figure 3
Task Interest by Actualization Level and Context

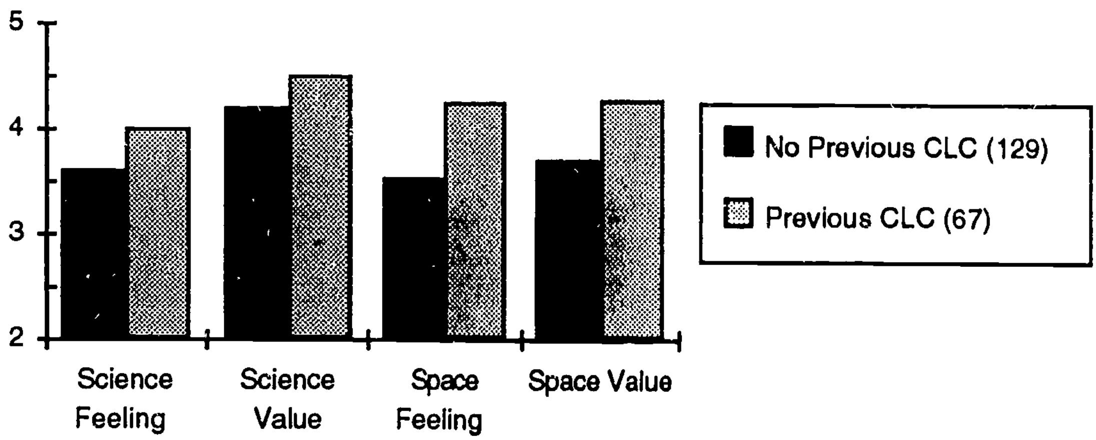
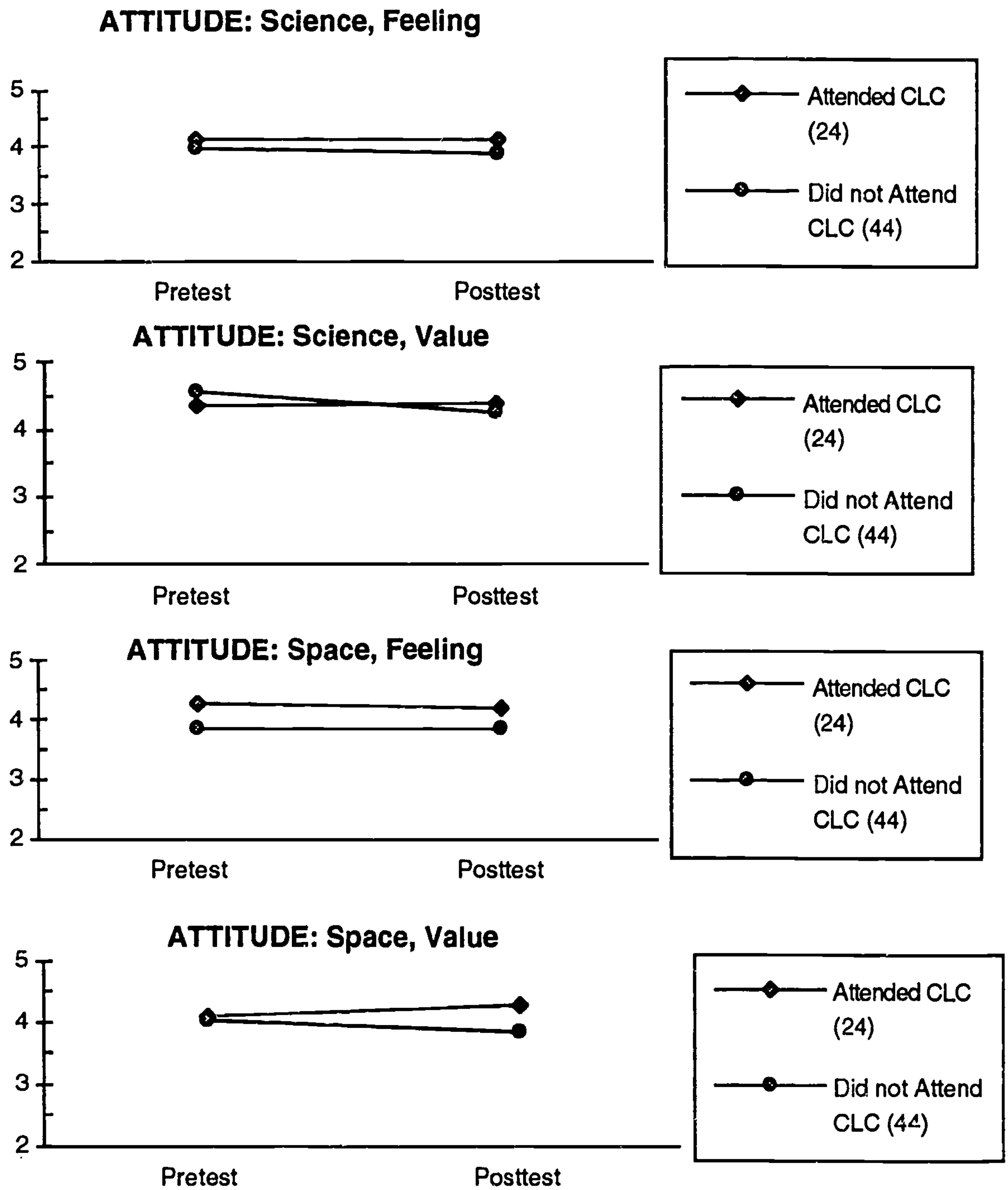
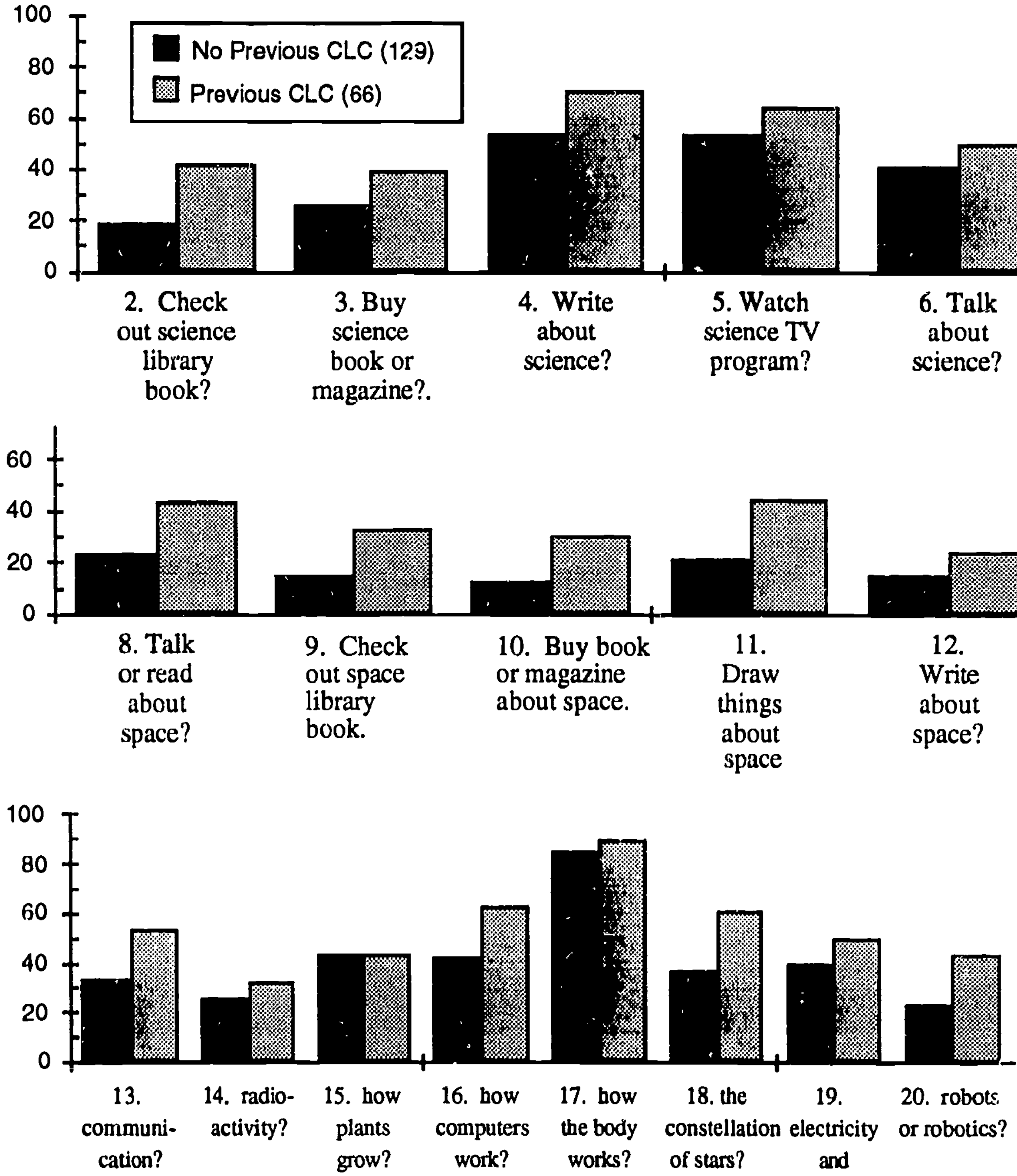


Figure 4.
Attitudes in Non-previous CLC/Previous-CLC Students



Figures 5a thru 5d
Site 2 Pretest-Posttest Attitude Change



Figures 6a, 6b, and 6c
Percentage of Site 1 Students Participating in Science-Related, Space-Related, and M.A.R.S.-Related Activities at Pretest

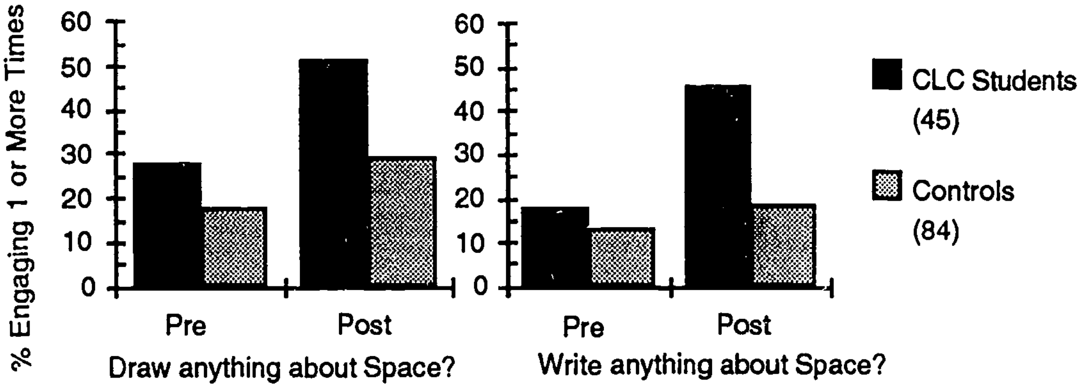


Figure 7
Site 1 Space-Related Activities at Pretest and Posttest

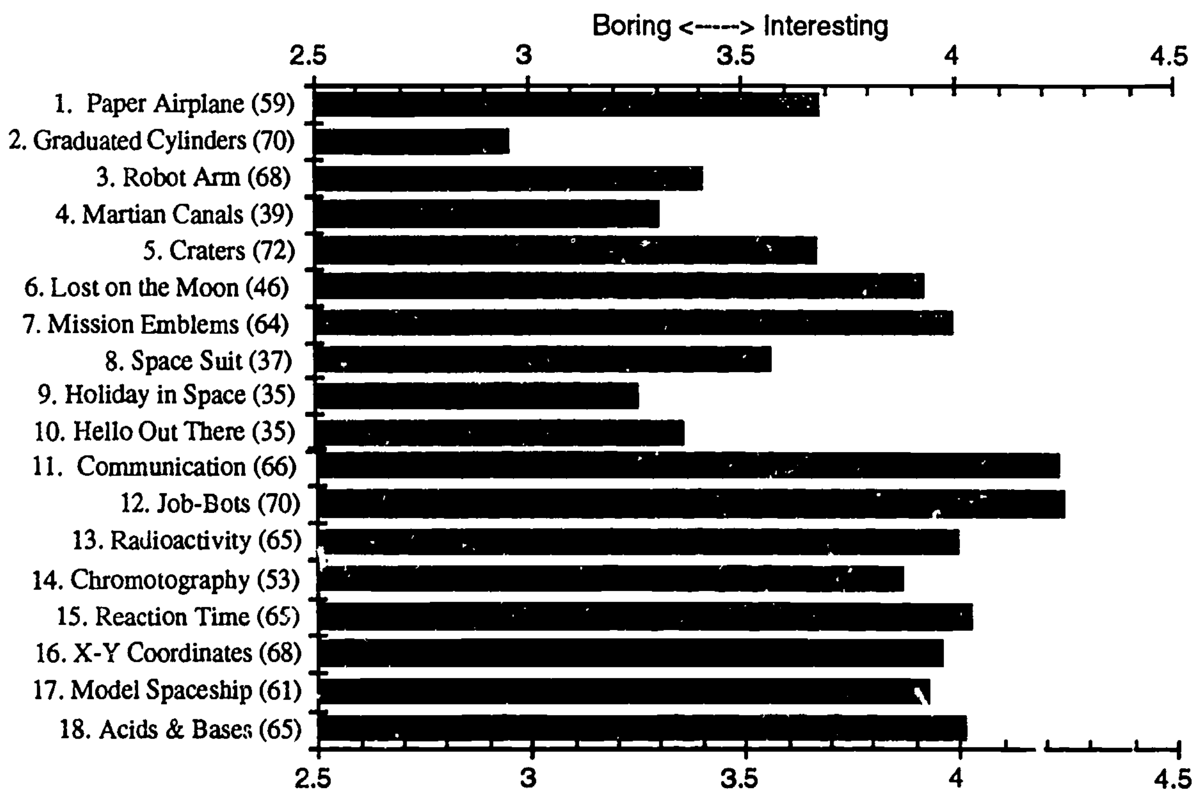


Figure 8
Interestingness of Preflight Classroom Activities.

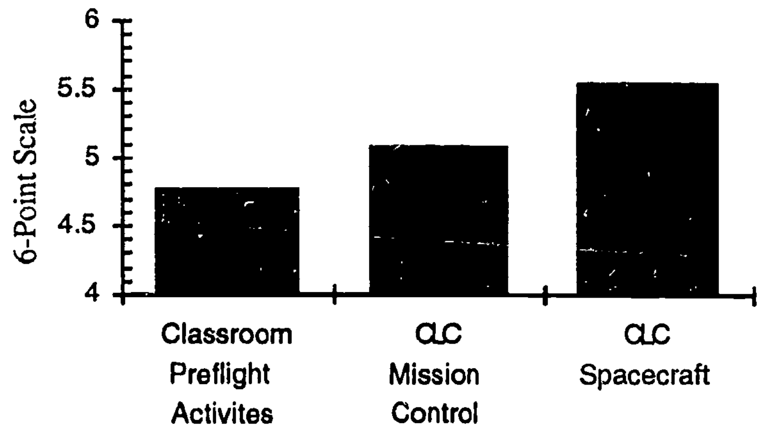


Figure 9
Interestingness of Major M.A.R.S. Components

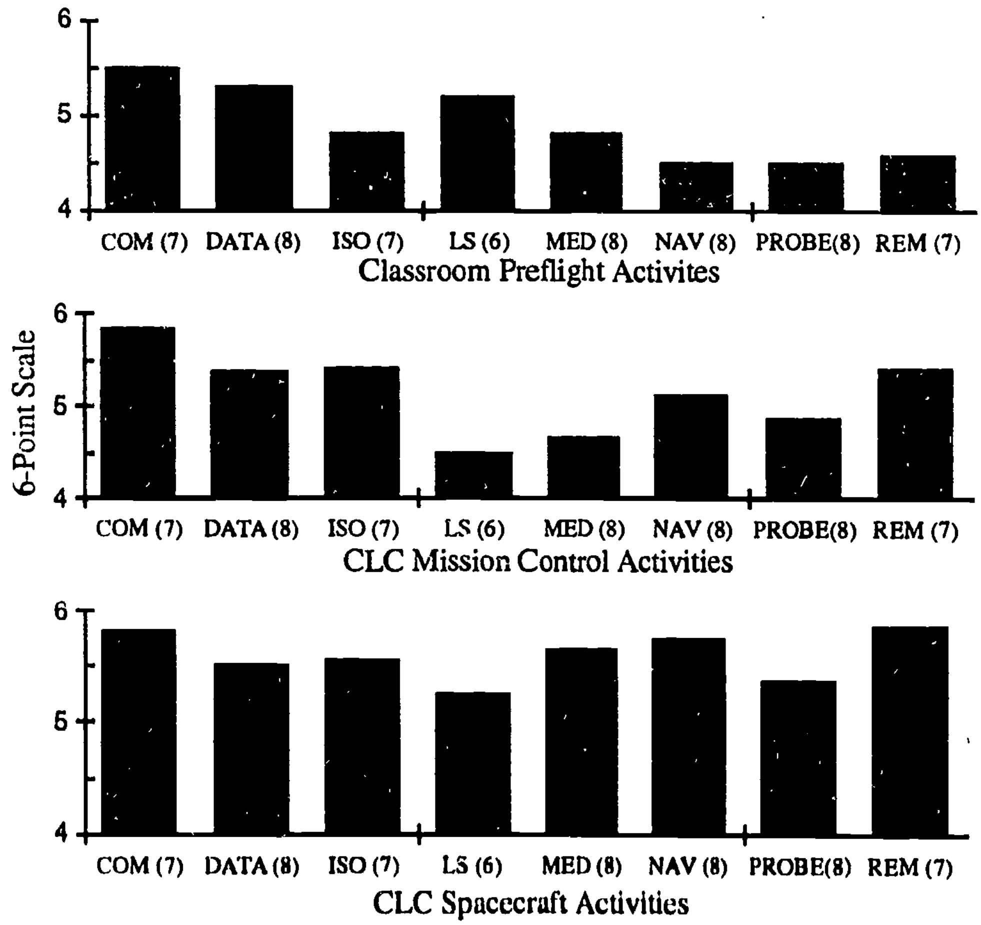


Figure 10
Interestingness of M.A.R.S. activities by team

Appendix A

SPACE/SCIENCE TAKS INTEREST SURVEY ITEMS

(PHYSICAL/Machines & Work)
Learning about machines and how they are used to do work?
Reading about how an motorized vehicle works?
Reading about how a Mars robot vehicle works?
Figuring out the forces that affect how a motorized wheelchair works?
Figuring out the forces that affect how a Mars rover vehicle works?
(PHYSICAL/Electricity and Magnetism)
Learning about electricity?
Reading about how electricity is produced and used?
Reading about how electricity is produced and used a spacecraft?
Figuring which is the most effecient way to make electricity?
Figuring out the best way to make electricity in a spacecraft?
(PHYSICAL/Sound, Heat, & Light)
Learning about sound, heat, and light?
Reading about how light and sound travel?
Reading about how light and sound travel in space?
Figuring out how long it takes light to travel one mile?
Figuring out how long it takes light to travel from the Sun to Mars
(LIFE/Animals)
Learning about animals?
Reading about how certain animals develop and reproduce?
Reading about how certain animals develop and reproduce in weightlessness?
Figuring out the best animals to use in a self-contained terrarium?
Figuring out the best animals to use in self-contained colony on Mars?
(LIFE/Plants)
Learning about plants?
Reading about how certain plants grow in greenhouses?
Reading about how certain plants grow in the weightlessness of space?
Figuring out how many trees are needed to make enough oxygen for one person?
Figuring out how many trees are needed to make enough oxygen for a Mars colony
(LIFE/Human Body)
Learning about the human body?
Reading about how age affects how the human body works?
Reading about how weightlessness in space how the human body were?
Figuring out which food provides the best nutrition for a camping trip?
Figuring out which food provides the best nutrition for a two-year space mission?
(EARTH & SPACE/Astronomy)
Learning about the solar system and the universe?
Reading about the different things orbiting the Earth?
Reading about the different things orbiting the Sun?
Figuring out how long it should take to travel to the Moon?
Figuring out how long it should take to travel to Mars?
(EARTH & SPACE/Geology)
Learning about Geology
Reading about the geological forces which created the Rocky Mountains?
Reading about the how the Valles Marineras (Mariner Valley) on Mars was formed?
Figuring out how the Great Smokey Mountains could have been formed?
Figuring out how Olympus Mons (a mountain on Mars) could have been formed?