ornl

ORNL/TM-10109

OAK RIDGE
NATIONAL
LABORATORY

MARTIN MARIETTA

Final Report: Manipulator Comparative Testing Program

J. V. Draper
S. J. Handel
E. Sundstrom
J. N. Herndon
Y. Fujita
M. Maeda

CONSOLIDATED FUEL REPROCESSING PROGRAM

OPERATED BY
MARTIN MARIETTA ENERGY SYSTEMS, INC.
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

Printed in the United States of America. Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road, Springfield, Virginia 22161
NTIS price codes—Printed Copy: A05 Microfiche A01

ORNL/TM-10109
Category UC-86

Consolidated Fuel Reprocessing Program

FINAL REPORT: MANIPULATOR COMPARATIVE TESTING PROGRAM

J. V. Draper, S. J. Handel, and E. Sundstrom
Clarke Ambrose, Inc.
10804 Sonja Drive
Knoxville, TN 37922-1064

J. N. Herndon
Fuel Recycle Division

Y. Fujita and M. Maeda
Power Reactor and Nuclear Fuels Development Corporation
PNC Tokai Works
Muramatsu, Tokai-mura
Nata-gun, Ibaraki-ken 319-11 Japan

Date Published: February 1987

Prepared by
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37831
operated by
MARTIN MARIETTA ENERGY SYSTEMS, INC.
for the
U.S. DEPARTMENT OF ENERGY
under Contract No. DE-AC05-84OR21400

CONTENTS

ABSTRACT

The Manipulator Comparative Testing Program compared the performance of selected remote manipulator systems under typical remote handling conditions. Three experiments were conducted to examine differences among three manipulator systems from the United States and Japan in the performance of typical remote handling tasks. The manipulator systems included the Meidensha BILARM 83A, the Central Research Laboratories Model M-2, and the GCA PaR Systems Model 6000. Six manipulator and control mode combinations were evaluated: (1) the BILARM in master/slave mode without force reflection, (2) the BILARM in master/slave mode with force reflection, (3) the Model M-2 in master/slave mode without force reflection, (4) the Model M-2 in master/slave mode with force reflection, (5) the BILARM with switchbox controls, and (6) the PaR 6000 with switchbox controls. The experiments examined differences between master/slave systems with and without force reflection and differences between master/slave systems and switchbox-controlled systems. A fourth experiment examined the relative contributions of the remote viewing system and the manipulator system to the performance of remote handling tasks.

Results of the experiments showed that operators using the Model M-2 in master/slave mode had significantly faster times to completion than operators using the BILARM in master/slave mode, with approximately the same error rate per trial. Operators were slower using the BILARM with force reflection than without it, and they committed more errors. There was no statistically significant difference between force-reflection and nonforce-reflection conditions for the M-2 manipulator for any of the performance criteria. However, these experiments were designed to test a wide range of manipulator systems. Tasks and procedures used in this testing were not sensitive to differences within any single system. No inferences about the effect of force reflection on remote task performance should be made from these data. A study designed to investigate the effect of various levels of force reflection on operator performance for a single manipulator system is presently ongoing separately at ORNL. The two manipulator systems in switchbox mode had significantly slower times to completion than any system in master/slave mode, with approximately the same error rate per trial. There were no significant differences between the BILARM in switchbox mode and the PaR arm. On average, tasks performed manually with television viewing took twice as long to perform as with direct viewing.

1. THE COMPARATIVE TESTING PROGRAM

1.1 PURPOSES

The general purpose of the Manipulator Comparative Testing Program was to evaluate and document the performance of selected manipulator systems from the United States and Japan. The manipulator systems were compared in the context of completion of typical tasks that might be encountered in the remote maintenance of nuclear fuel recycling facilities of the future. The Meidensha BILARM 83A (Japan), the Central Research Laboratories (CRL) Model M-2 (U.S.), and the GCA/PaR Systems Model 6000 (U.S.) were tested in this program.

The Manipulator Comparative Testing Program was jointly sponsored by Martin Marietta Energy Systems, Inc., (for the U.S. Department of Energy) and the Power Reactor and Nuclear Fuel Development Corporation (PNC) of Japan. The site of the testing was the Remote Operations and Maintenance Development (ROMD) facility, which is operated by the Fuel Recycle Division of Oak Ridge National Laboratory (ORNL). The facility consists of a high-bay remote handling equipment demonstration area that contains prototypical process equipment and manipulator systems, along with a control room for the manipulators. Figure 1 shows the high-bay area, and Fig. 2 shows the control room.

ORNL–PHOTO 4592–84

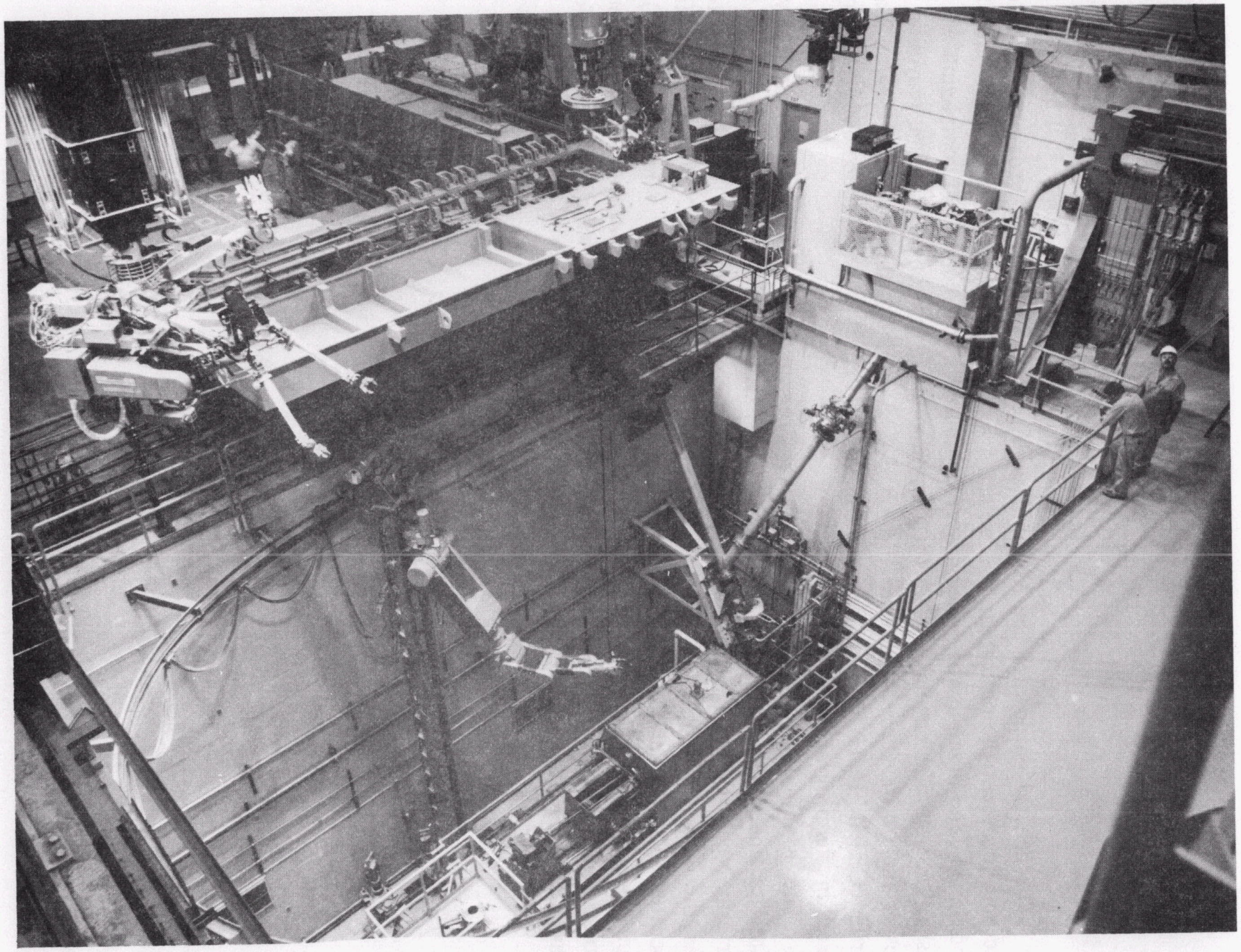

Fig. 1. The high-bay area at the ROMD facility.

ORNL–PHOTO 8739–84

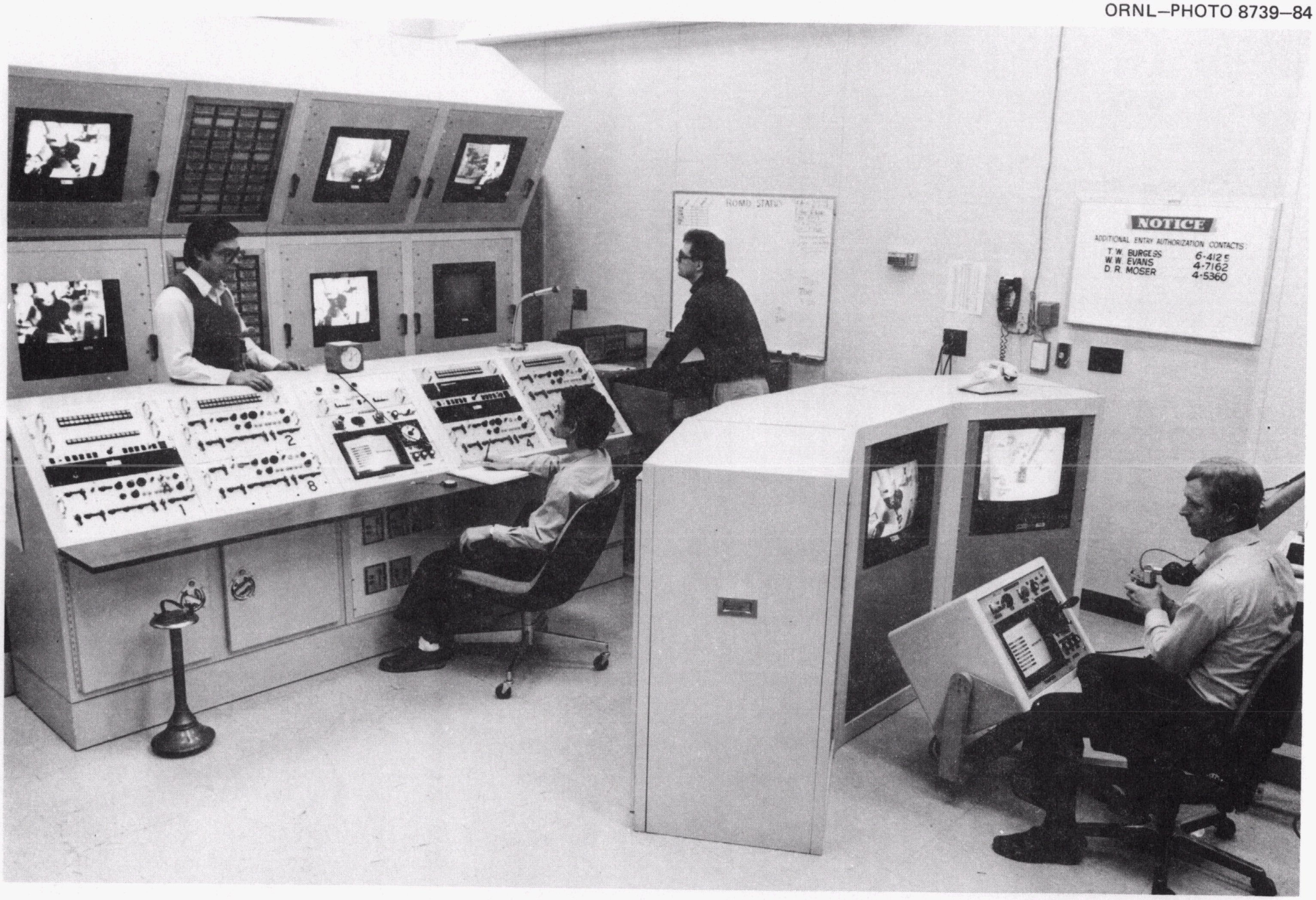

Fig. 2. The ROMD facility control room.

2. MANIPULATOR SYSTEMS

2.1 THE BILARM 83A

The Meidensha BILARM 83A is a unique, force-reflecting master arm-controlled electromechanical manipulator system. It is unique in that the slave arm can also be operated as a switchbox-controlled manipulator (with no force feedback) using a position control switchbox. Figure 3 shows the master arm, and Fig. 4 shows the switchbox controller. Figure 5 shows the slave arm, which has a handling capacity of 25 kg in any position and 8 D.F. (degrees of freedom), including the gripper. Table 1 lists the range of motion and speed for each of its joints.

The BILARM 83A slave arm joints are each driven by a packaged drive unit consisting of a permanent magnet dc motor, planet gear drive and harmonic gear drive in series, and a plastic film potentiometer for position signals. The torque at the slave arm is measured at each of three joints by a load cell, which is located to minimize friction losses. The force signal is used as a drive signal for the three force-reflecting master joints and as local feedback to the slave. Force-reflection ratios (the ratio of force applied at the slave to force output at the master arm) from 1:1 to 9:1 are available, as well ∞:1 (no force reflection). The switchbox control is unique in that it uses position controls instead of rate controls usually applied to these types of devices. The potentiometers on the BILARM 83A switchbox directly control slave-joint positions, rather than joint velocity.

ORNL–PHOTO 8737–84

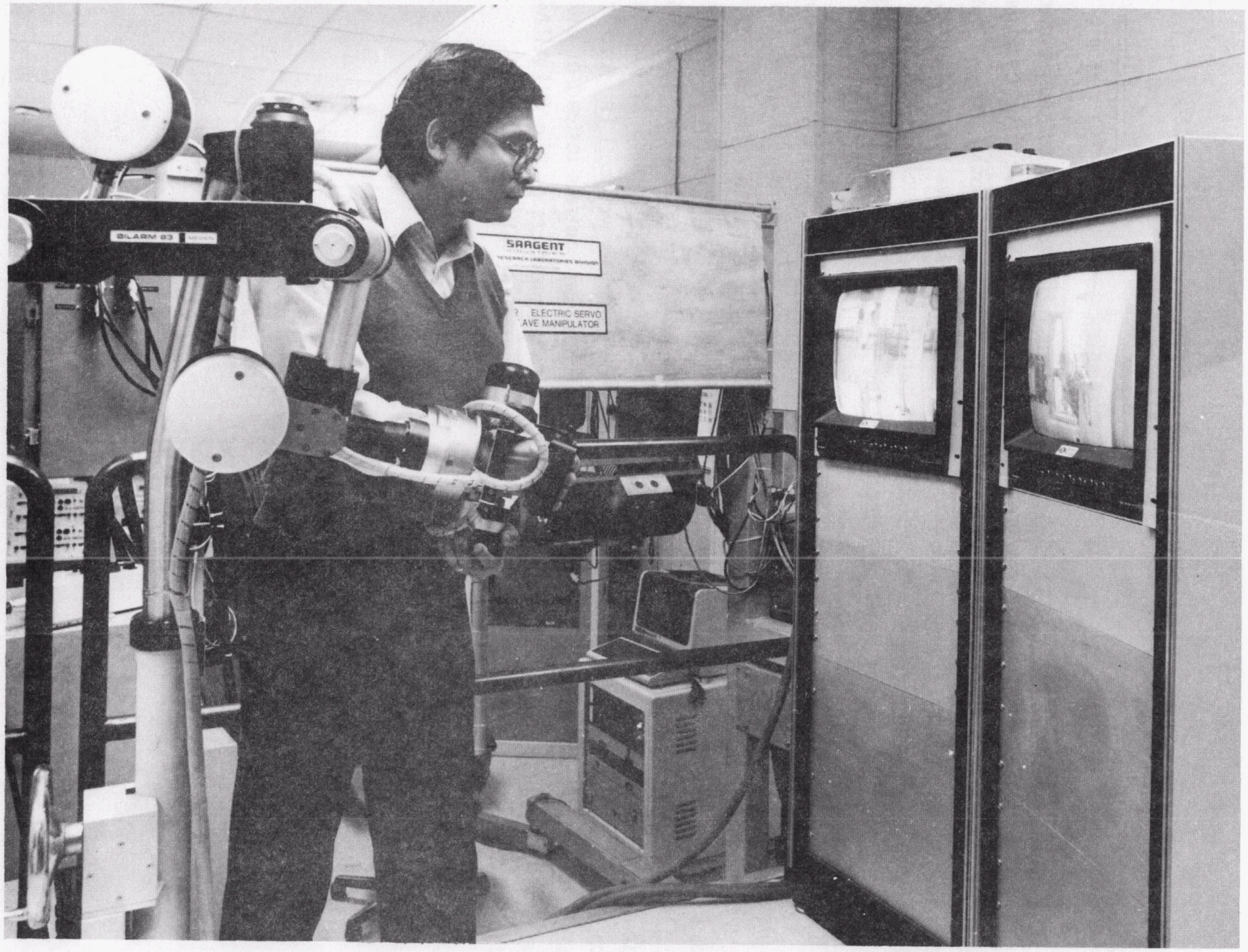

Fig. 3. The BILARM 83A master controller.

ORNL—PHOTO 8738—84

Fig. 4. The BILARM 83A switchbox controller.

ORNL–PHOTO 2488–85

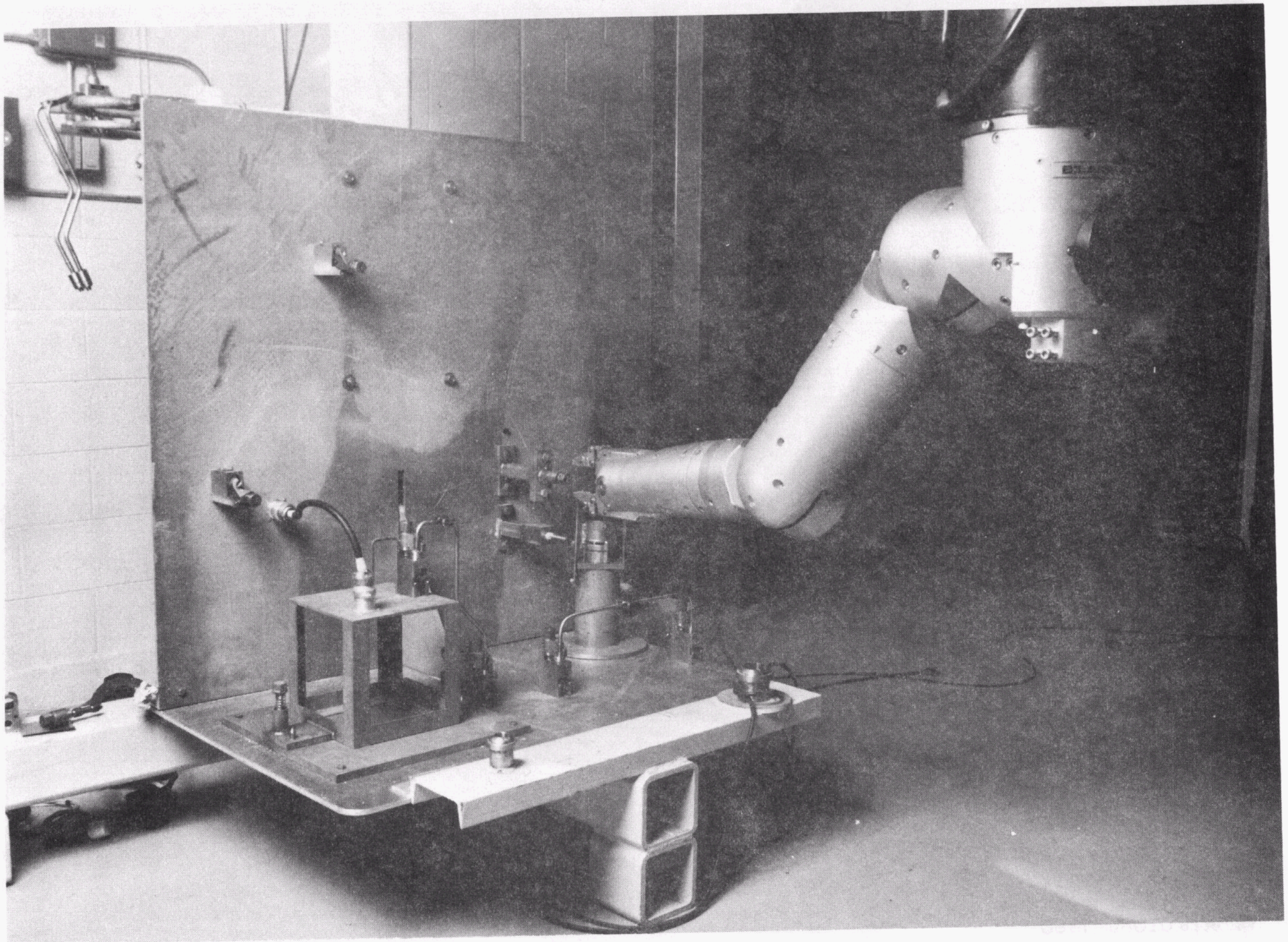

Fig. 5. The BILARM 83A slave arm.

Table 1. Ranges of motion and speed of BILARM 83A joints

Joint	Range of motion[a]	Maximum no-load speed
Shoulder rotation	±200°(±45°)	20°/s
Shoulder swing A (left-down-right)	±80°(±45°)	8°/s
Shoulder swing B (down-front-up)	±82°	8°/s
Elbow rotation	±75°	10°/s
Elbow swing	±67°	10°/s
Wrist rotation	±115°(±110°)	50°/s
Wrist swing	±90°(±45°)	50°/s
Finger grasping	120 mm	18.5 mm/s

[a]Ranges in parentheses controlled by master arm.

2.2 THE CRL MODEL M-2

The CRL Model M-2 manipulator is a bilateral, force-reflecting servomanipulator system. The master arms (Fig. 6) are 7-D.F. kinematic replica controllers. The slave arms (Fig. 7) each have a handling capacity of 23 kg continuous in any position. The kinematics are in the typical "elbows-up" stance used by all of the currently commercially available servomanipulators in the United States and Europe. Table 2 lists the range of motion and speed for each joint.

The M-2 slave joints are each driven by a brushless dc servomotor with integral position and velocity encoding. The outputs of the three upper degrees of freedom are gear and lever driven. The four lower degrees of freedom of master and slave are cable driven. The master controller lower degrees of freedom are tape driven. A standard position–position technique, implemented in digital control hardware and software, provides force reflection. Force-reflection ratios from 1:1 to 8:1 are available, as well as ∞:1 (no force reflection).

ORNL—PHOTO 8736–84

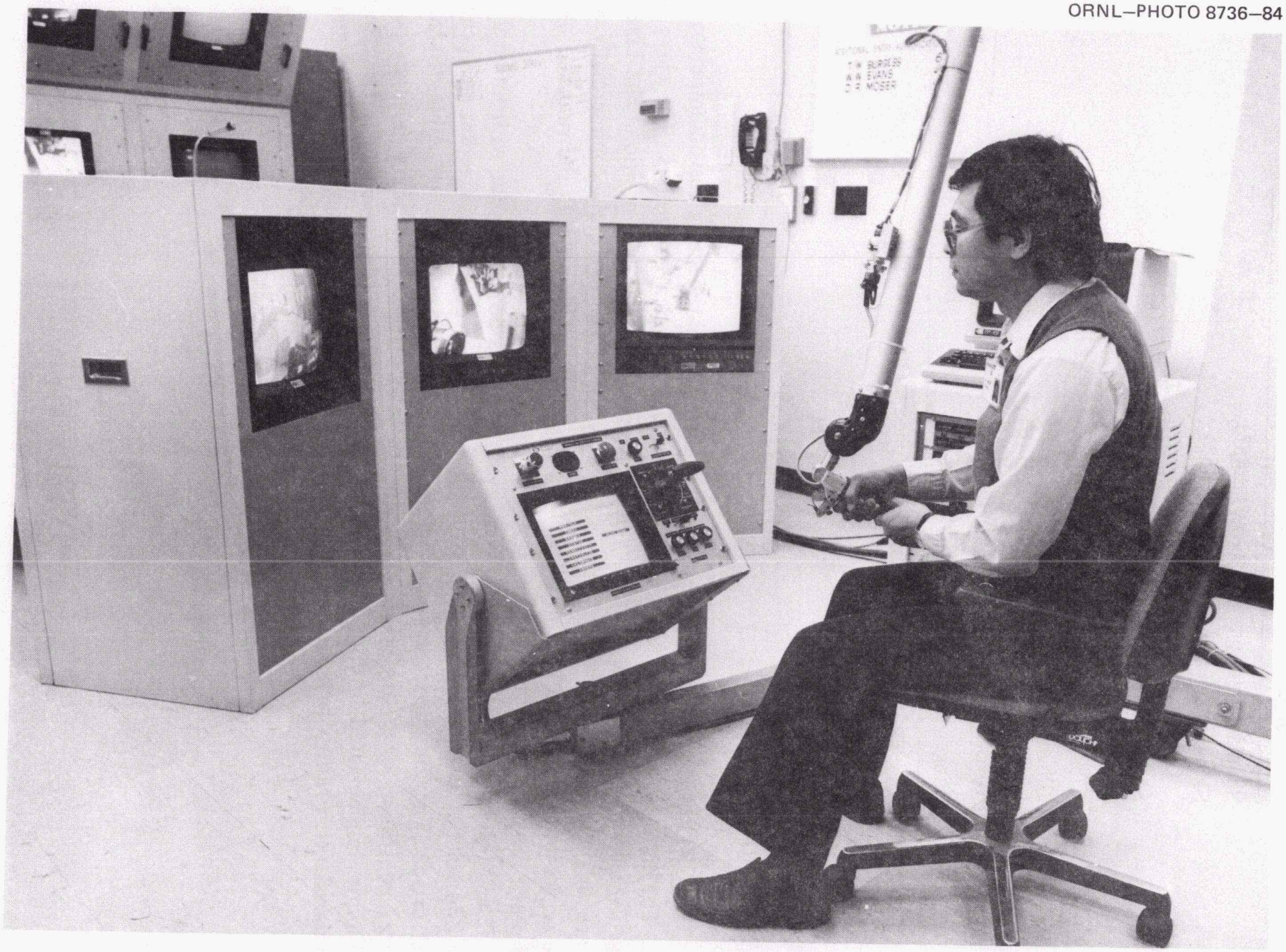

Fig. 6. The CRL Model M-2 master controller.

ORNL—PHOTO 2487—85

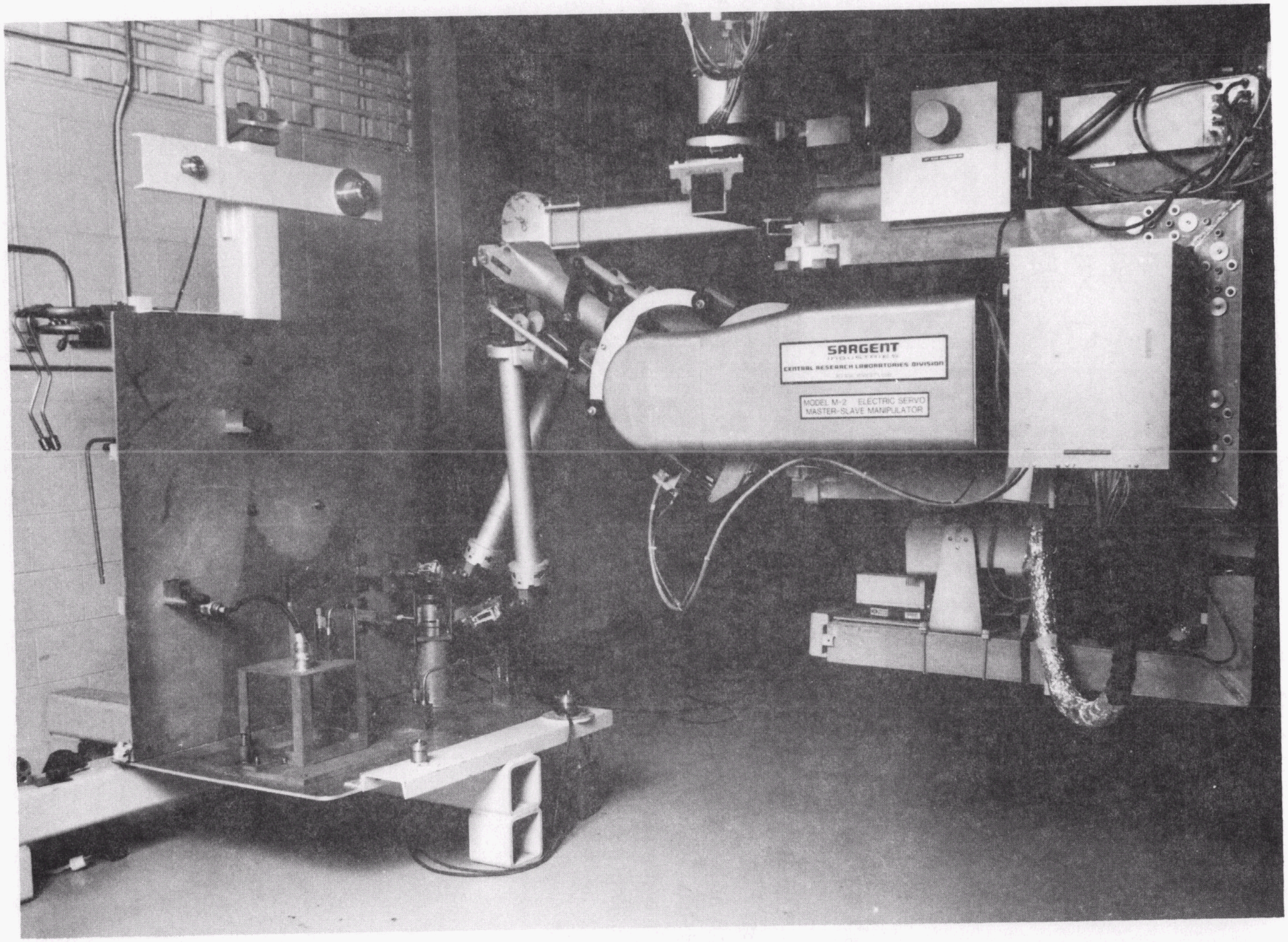

Fig. 7. The CRL Model M-2 slave arm.

Table 2. Ranges of motion and speed of CRL Model M-2 joints

Joint	Range of motion[a]	Maximum no-load speed
Shoulder roll	±45°	>1.5 m/s
Elbow pitch	±45°	>1.5 m/s
Shoulder pitch	±45° (255°)[a]	>1.5 m/s
Wrist yaw	±210°	>344°/s
Wrist patch	±40°; −125°	>400°/s
Wrist rotation	±180°	>344°/s
Gripper closure	Gripper, 0.08 m Handle, 0.07 m	>1 m/s

[a]Range in parentheses equivalent to total with indexing.

2.3 THE PaR MODEL 6000

The PaR Model 6000 is a rate-controlled power manipulator of typical design for its class. Figure 8 shows the switchbox controller. The controller provides rate control of the slave arm, which is shown in Fig. 9. The slave arm has a 181-kg capacity in all positions and has 7 D.F., including the gripper. The slave arm uses permanent magnet dc motors with continuously variable speed control. Table 3 lists the range of motion and speed of each joint.

ORNL–PHOTO 6838–85

Fig. 8. The PaR Model 6000 switchbox controller.

ORNL—PHOTO 0840–85

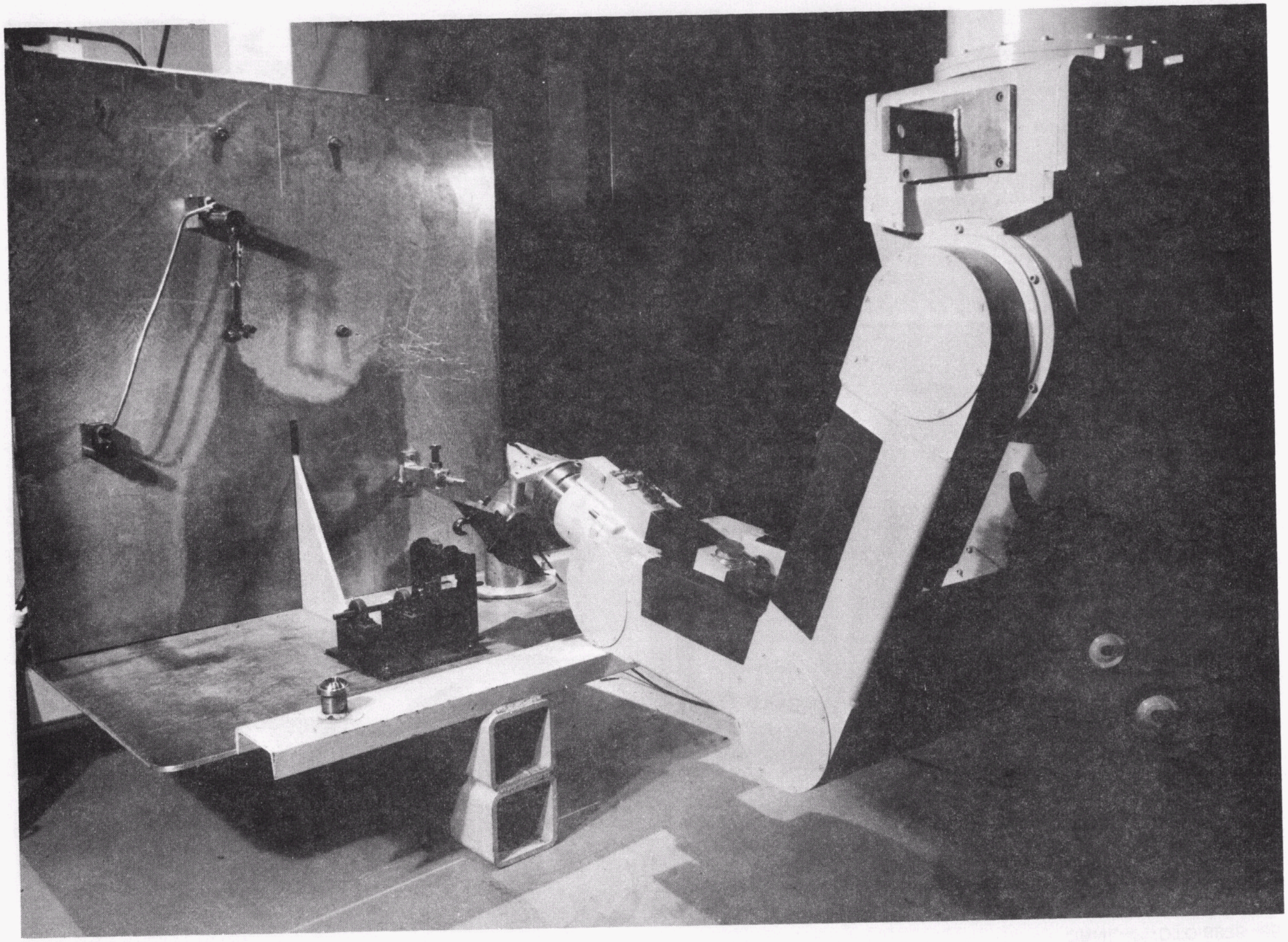

Fig. 9. The PaR Model 6000 slave arm.

Table 3. Ranges of motion and speed of PaR 6000 joints

Joint	Range of motion[a]	Maximum no-load speed
Shoulder rotation	$\pm\infty$	1.0 rpm
Shoulder pivot	±105°	1.0 rpm
Elbow pivot	±120°	1.0 rpm
Wrist pivot	±150°	1.0 rpm
Wrist extension	152 mm	0.5 m/min
Wrist rotation	$\pm\infty$	7.0 rpm
Gripper closure	203 mm	0.4 m/min

3. PLAN OF RESEARCH

This section describes the research objectives, the design of the research, the general procedures used in collecting the data, and the strategy for analyzing the data.

3.1. RESEARCH OBJECTIVES

The Manipulator Comparative Testing Program examined the performance of the Japanese manipulator system (BILARM 83A) and the U.S. manipulator systems (CRL M-2 and PaR 6000) when they were used to complete typical remote handling tasks. The program primarily sought to identify and document differences between the systems. The program also investigated (1) whether the presence of force reflection affected the performance of manipulator systems in the master/slave mode, and (2) whether performance of the BILARM with master/slave control differed from its performance with switchbox control. As an adjunct to the central research topics, the program examined the relative contribution of the viewing system and the manipulator system to overall performance by comparing performance with the manipulator systems to direct-contact maintenance with direct and television viewing.

3.2 RESEARCH DESIGN

The program included four interrelated experiments. The strategy of the testing was to begin with very simple tasks under tightly controlled conditions and to include progressively more realistic tasks and conditions as the testing progressed. In experiment 1, five operators repeatedly completed a set of three simple tasks, using each of the manipulator systems. Performance criteria included the time required to complete each task (time to completion), the number of errors committed during completion of each task, and the time required to complete each subtask of the experimental task. The results of the first experiment guided the development of a testing plan for experiment 2. A slightly more complex set of tasks was used, and, as in experiment 1, the performance criteria included time to completion and time allocated to subtasks. However, the error measure was expanded to include a separate tallying of 18 types of critical incidents, that is, those events that may be important in the performance of remote handling tasks. These included not only errors but other important occurrences such as the use of two hands to control one manipulator arm. The operators also rated the manipulators on difficulty of use by completing a simple questionnaire in experiment 2.

In the third experiment, operators did a single, realistic task under conditions designed to simulate those that would be encountered in remote maintenance. Performance criteria were limited to task completion time and the total of all critical incidents because of the low number of trials completed, which made the individual critical incidents and subtasks statistically unstable.

Experiment 4 ran concurrently with the other three experiments. In this experiment, the operators performed experimental tasks manually using direct viewing and television viewing. The results permitted evaluation of the relative importance of the viewing system to the performance of the remote handling system.

3.3 RESEARCH FACILITY

The Remote Operations and Maintenance Demonstration (ROMD) facility provided a realistic environment for remote maintenance testing. The ROMD facility consists of a large high bay for the simulated remote handling area and an adjoining control room. The controls of the manipulator systems ("master" controller and "switchbox" components) for these experiments were located in the control room, and the slave manipulators were located in the high-bay area. The slave components of the manipulator systems were positioned in front of an aluminum panel and shelf on which the tasks were mounted (the task board). In each room a video camera was mounted for recording the experimental sessions. A control room camera was mounted and focused to give a sideview of the operator (from the waist up) and the controls of the manipulator system. In the high-bay area, the video camera was mounted and focused to provide a side view of the slave portion of the manipulator system and the task board.

Conditions in the testing rooms were prescribed in written procedures prepared prior to each stage of experimentation. An observer was required to give instructions to the operators and to record the times to completion of the tasks and the errors committed during testing. A second observer monitored the video cameras and associated equipment and made videotape recordings of the sessions.

3.4 GENERAL PROCEDURE

Each experiment included six combinations of manipulator systems and modes of control: (1) the BILARM with force reflection, (2) the BILARM without force reflection, (3) the Model M-2 with force reflection, (4) the Model M-2 without force reflection, (5) the BILARM with position control switchbox, and (6) the PaR 6000 with rate control switchbox. A team of trained operators completed a set of remote handling tasks with each of the systems under strictly controlled conditions. An observer recorded the amount of time taken to complete each task (time to completion) and a measure of the quality of performance. The measure of performance quality was either the number of errors committed (experiment 1) or the number of times each of a set of 18 critical incidents occurred (experiments 2 and 3). A second observer supervised the videotape recording of each session.

Operators who participated in the program were (with one exception) drawn from a pool of experienced manipulator operators available at ORNL's Fuel Recycle Division. The exception was a Japanese engineer, who was responsible for installing and maintaining the

BILARM and was the most experienced operator available for that system. Three of the operators participated in all four experiments; two others participated in experiment 1 and part of experiment 4. One other operator participated in experiments 2, 3, and 4.

Operator performance was expected to improve with practice. In order to ensure that improvement in performance with practice did not bias results relative to the primary research questions, each operator was presented with a different order in which to use the manipulators. This equalized the amount of practice with each manipulator.

All sessions were videotaped. Each session was recorded on a video cassette, using two video cameras and a video recorder. All recordings were split-screen images, with half of the screen devoted to the operator and master controls and half devoted to the slave components and the task.

3.5 STRATEGY OF ANALYSIS

The general strategy of analysis was to compare the average values of the criteria among the six combinations of manipulator and mode. The computational procedures (derived from O'Brien and Kaiser[1]) were specifically designed for the situation in which each of the set of operators was tested under all of the combinations of machine and mode of control. The statistical analyses incorporated a specially adapted version of the analysis of variance (ANOVA). Each comparison was based on a separate constrast between averages of the dependent measures calculated for two or more of the experimental conditions. For example, in comparing the BILARM with the M-2 on time to completion of the impact-wrench task, the average of each operator's scores from sessions involving the BILARM was contrasted with the average from sessions involving the M-2. The average of differences for the operator group was tested for significance using a dependent t-test[2] with $n - 1$ D.F., where n was the number of operators. A t-test was considered significant if it exceeded the criterion value of t at $\alpha \leq 0.05$. In this study, t is reported for significant tests along with the minimum value of α for the test. The appendix contains a detailed discussion of the t-test and of alpha.

4. EXPERIMENT 1

In experiment 1, five operators performed three simple tasks. The best performance occurred with the M-2 manipulator; operators completed tasks in the shortest time with the fewest errors. The second best performance occurred with the BILARM in master/slave mode. There was no difference between the BILARM in switchbox mode and the PaR manipulator, and there was no difference between force-reflection conditions.

4.1 TASK DESCRIPTIONS

The three tasks selected for experiment 1 were designed to be representative of remote handling tasks primitives for fuel recycle facilities and other remote handling applications. The tasks were also designed so that they could be easily performed with each of the combinations of machine and mode of control and could be completed several times in a 1-h session. Figure 10 shows the three tasks in experiment 1. The tasks are described in detail in the following sections.

ORNL–PHOTO 8745–84

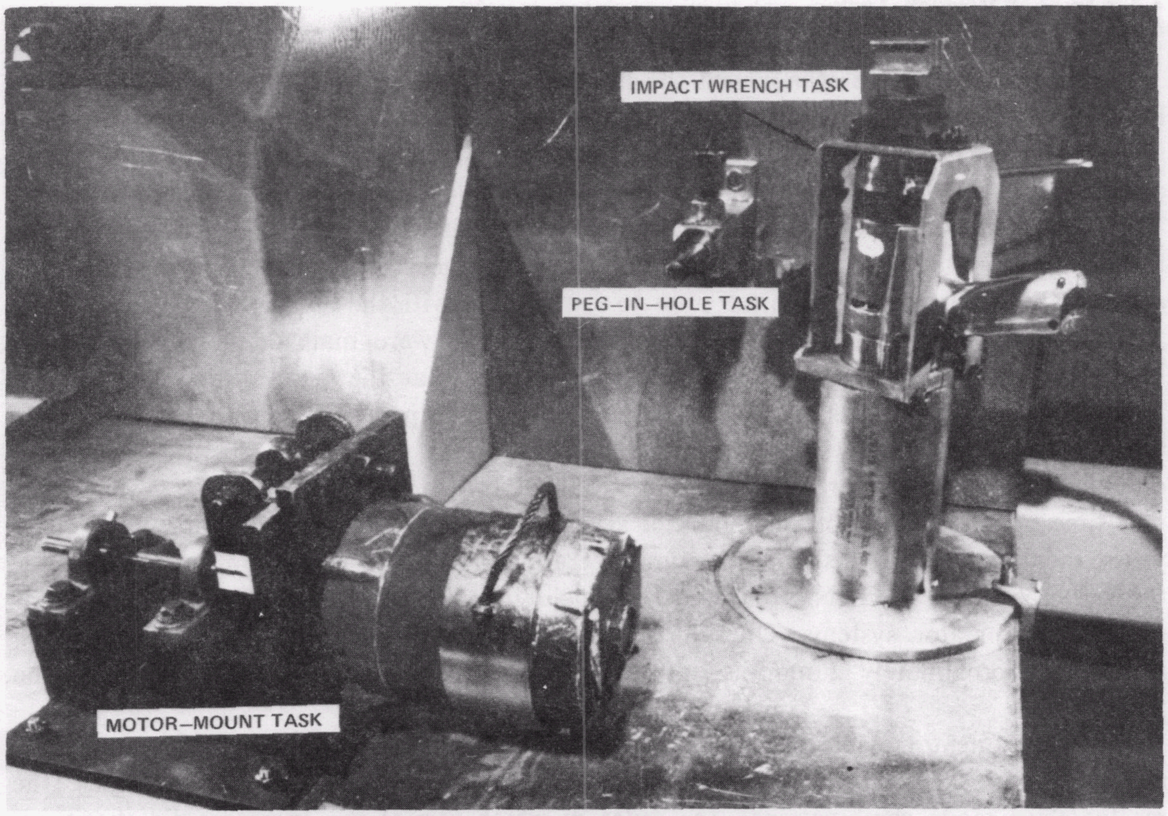

Fig. 10. Tasks used in experiment 1.

4.1.1 Motor-Mount Task

In this task the operator moved the manipulator end effector from its starting position to the location of an electrical motor, which was mounted on a horizontal metal rod by means of a hook. The hook and rod arrangement facilitates remote connection of the motor with a geared shaft attached to the mount. The operator used the manipulator to grasp a steel cable attached to the motor, lift the motor, and disengage the hook from the mount. After the motor was disengaged, the operator moved it to touch a predesignated position on the task board shelf. The motor was then returned to the location of the motor mount, and the hook was reengaged onto the rod. The task was completed when the manipulator arm returned to its starting position.

4.1.2 Peg-in-Hole Task

In this task the operator moved the manipulator from its starting position to the location of a cylindrical peg, grasped the peg, extracted it from the hole in which it was mounted, moved it to touch a predesignated area on the task board, returned it to its hole, and inserted it to its full depth. The peg was made of stainless steel and was 0.742 in. in diameter and 6 in. long. It was mounted in a horizontal hole, 0.750 in. in diameter, with its end facing the manipulator arm. The peg had two flat faces cut onto opposite sides to form a handle and a slight taper at the end to ease insertion.

4.1.3 Impact-Wrench Task

In this task the operator moved the manipulator from its starting position, grasped a bracket attached to a mechanical impact wrench, lifted the wrench from its cradle, moved the wrench to a position immediately above a hexagonal bolt that protruded vertically from a mount on the task board, and lowered the wrench so that its socket engaged the bolt. The operator then removed the wrench, returned it to its cradle, released it, and returned the manipulator to its starting position. The impact wrench socket was oversized; therefore, angular alignment was not required prior to engagement with the bolt.

Five operators participated in experiment 1. Four were male employees of Martin Marietta Energy Systems who were working at ORNL at the time of the Comparative Testing Program. One was a male employee of PNC in Japan. Operator 1 had worked at ORNL for approximately 11 years, including 5 years as a part-time operator of remote manipulator systems. Operator 2 had worked at ORNL for 9 years, including 2 years as a part-time operator of remote manipulator systems. Operator 3 had worked at ORNL for 2 years, including 18 months as a part-time operator of remote manipulator systems. Operator 4 had worked at ORNL for 34 years, including 30 years as a full-time operator of remote manipulator systems. Operator 5 had worked for PNC for a total of 5 years, including approximately 2 months during which he occasionally practiced operating the BILARM.

In experiment 1, each operator was assigned a unique sequence for using the manipulator systems. The sequences were designed so that for the first four combinations (the master/slave mode combinations), the order in which two of the operators did them was exactly the reverse of the order in which another operator did them. The remaining operator (No. 2) had an order in which the machines and the presence of force reflection were alternated. Only the first four operators completed testing in the switchbox mode; operator 5 returned to Japan before the testing of manipulators in the switchbox mode was conducted. The order of testing of the two switchbox combinations was the same for all four operators because of the physical constraints of the facility. Switchbox combinations were always tested after the four master/slave combinations.

Table 4 lists the testing sequence for each of the five operators in experiment 1. Each operator completed two sessions with each manipulator before moving on to the next system in the series. A session consisted of five consecutive repetitions of each of the three tasks for a total of 15 trials. Each operator completed a total of 30 trials per machine/mode combination.

Table 4. Testing sequence for each operator in experiment 1

Operator	Sequence					
1	MF[a]	MN[b]	BF[c]	BN[d]	BS[e]	PS[f]
2	BN	MF	BF	MN	BS	PS
3	BF	MF	BN	MN	BS	PS
4	BN	BF	MN	MF	BS	PS
5	MN	BN	MF	BF		

[a]M-2 force reflecting.
[b]M-2 nonforce reflecting.
[c]BILARM force reflecting.
[d]BILARM nonforce reflecting.
[e]BILARM with switchbox.
[f]PaR manipulator.

4.2 PROCEDURES

Before beginning the testing trials in experiment 1, each of the four operators who were employees of ORNL were given 1 h of practice on the M-2 manipulator and 3 h of practice on the BILARM with master/slave controls. They were also given 1 h of practice with each of the two switchbox combinations before the trials involving switchbox controls. The operator who was an employee of PNC was relatively less experienced with the M-2 and relatively more experienced with the BILARM; therefore, he was given 3 h of practice with the M-2 and 1 h with the BILARM.

Each operator completed two sessions of trials under each of the six combinations of manipulator systems and modes—in the order indicated in the testing sequence. A session consisted of a set of 15 trials. Each set of 15 trials comprised five successive trials with one of the tasks, followed by five trials with a second task, and concluding with five trials with the third task. The order in which an operator performed the three tasks in a particular set of trials was determined in advance through the use of a table of random numbers. A rest period of at least 1 h followed each testing session.

An observer supervised each session to ensure that the operator was properly positioned at the controls of the manipulator system and that the slave components were in position for the task. The observer also ensured that the video cameras and videotape recorder were operating properly and had been started for the trial. At the observer's signal, the operator started the task, and the observer started a stopwatch. When the operator completed the trial, the observer stopped the stopwatch and the video recorder. The observer counted the errors committed during the trial and made a written record of task completion time and errors on a preprinted form.

4.3 DEPENDENT VARIABLES

The performance criteria used in experiment 1 included time to completion, number of errors, and time to completion of subtasks. The time-to-completion criterion consisted of the amount of time that elapsed between the signal to start a trial and its successful completion, excluding periods during which the trial was interrupted because of malfunction or the need to make a manual correction in the task area. For analysis of the total time to completion, the logarithm (base 10) of the score was used instead of the actual number of seconds elapsed. This is a standard procedure for analysis of duration data because conversion to logarithms prevents outliers (unusually high or low scores) from unduly influencing averages and variances.[3] Duration scores converted to logarithms are also more likely to be normally distributed than raw duration scores.

Errors were recorded by the test observer during the task. Occurrences of four types of errors were recorded: misalignments, misses, drops, and overloads/damage. Misalignments occurred when an operator attempted to position the manipulator or an object held in the grasp of the manipulator and failed to achieve the correct position. Examples included positioning the motor hook off-center, striking the surface of the peg-in-hole task with the peg instead of inserting it into the hole, and attempting to insert the peg into the hole with the peg off-center or not parallel with the axis of the hole. Groping movements that did not "break the plane" of the target were not counted as errors, such as when an operator attempted to insert the peg in the hold, moved it toward the hold, noticed that it was off-center, and corrected its alignment without first making contact. Misses included failures to grasp, pick up, or push an object, such as failure to grasp the cable attached to the motor in the motor-mount task. Drops included incidents when an item fell from the tongs of the manipulator. Overloads/damage included any overloads that activated the brakes, clutches, or alarms of the manipulator systems, or any events that caused visible damage to the manipulator or task area.

Each of the three tasks was analytically divided into an exhaustive series of sequential subtasks. The first and last trial for each of the three tasks in each session were analyzed by subtask. The sample included 6 of the 15 trials in each session, including 2 for each task. An observer measured the time taken to complete each subtask from the videotape recordings made of each session. Table 5 lists the subtasks for each task in experiment 1.

Table 5. Subtasks for the three tasks used in experiment 1

Subtask	Task: Impact wrench	Task: Peg in hole	Task: Motor mount
1	Move from start to wrench holder	Move from start to peg holder	Move from start to motor cable
2	Align tong and grasp wrench	Align tong and grasp peg	Align tong and grasp cable
3	Extract wrench and move to bolt	Extract peg	Disengage motor
4	Engage bolt with wrench	Move to panel, touch, and return	Move to panel, touch, and return
5	Move to wrench holder	Align peg with hole	Align motor hook and release
6	Insert wrench	Insert peg	

4.4 RESULTS

This section presents the results associated with each criterion for comparisons among the six combinations of manipulator system and mode of control. Results for the master/slave combinations are presented first, followed by the switchbox combinations, and then a comparison of master/slave and switchbox modes on the BILARM. Figure 11 illustrates the averages of task completion time, and Fig. 12 shows errors for the machine and mode combinations in experiment 1.

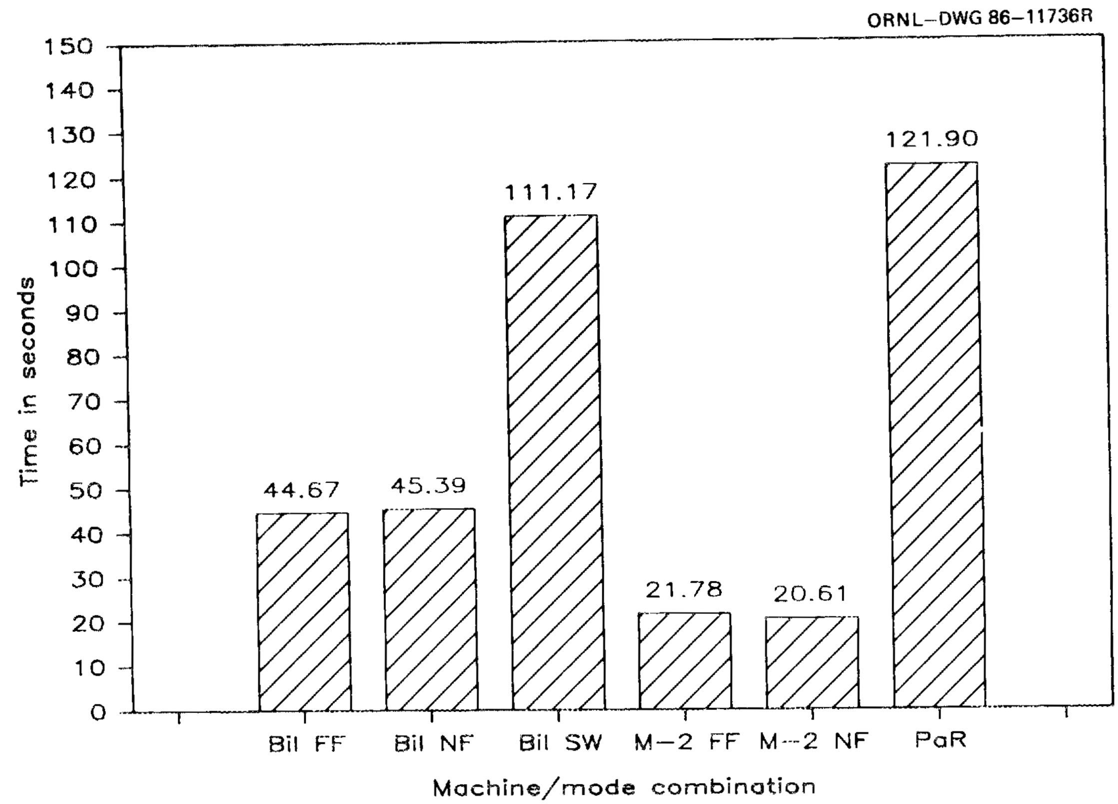

Fig. 11. Time to complete tasks.

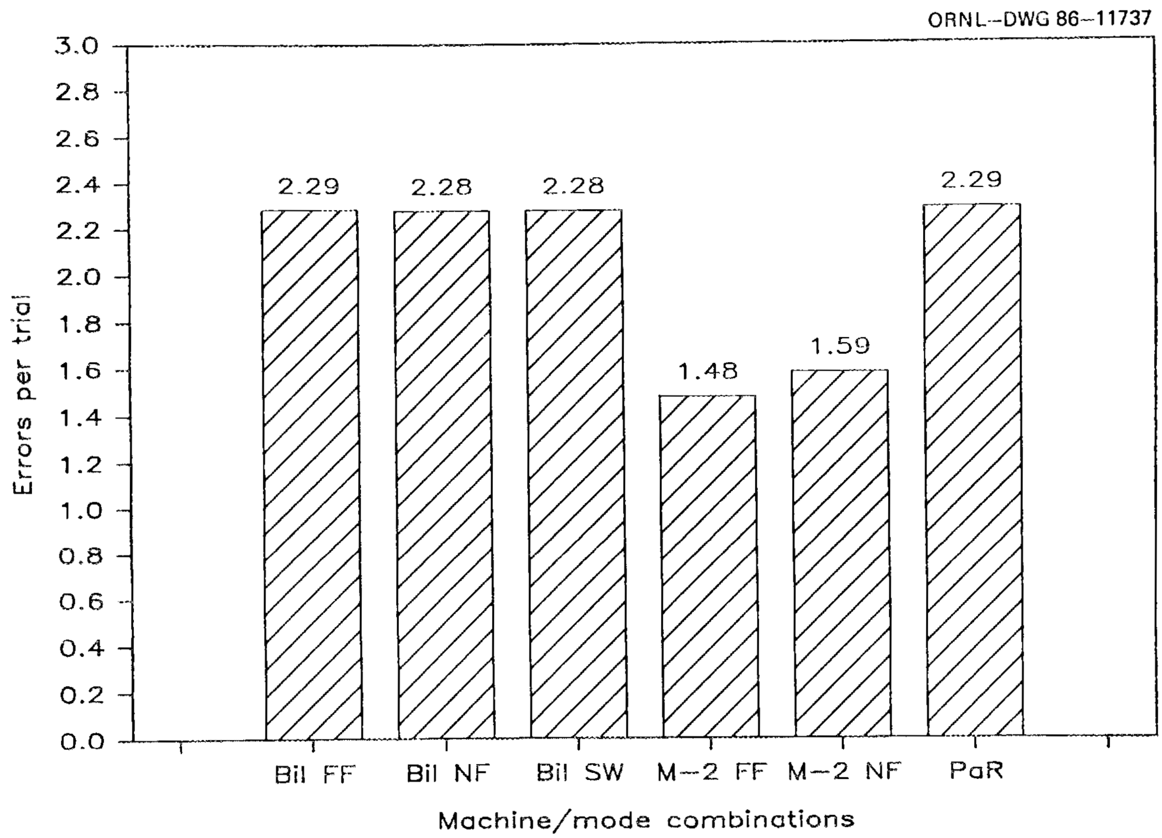

Fig. 12. Errors per trial.

4.4.1 Machine Differences Within Master/Slave Systems

Operators using the M-2 performed tasks significantly faster than they did with the BILARM. Table 6 shows the averages of task time and the logarithm to base 10 of task time for machines under master/slave control. Blank spaces under the t-test column (average errors) in Table 6 indicate nonsignificant t-scores. Five operators participated in the first experiment; therefore, all tests were conducted with 4 D.F. Table 6 shows that the averages of the logarithm to base 10 of times to completion for the BILARM were much longer than those for the M-2. When the averages of the logarithm to base 10 of task time were converted back to actual task time, the BILARM took twice as long as the M-2. The times to completion within tasks were also consistently shorter for the M-2.

The M-2 yielded fewer errors for each task, although the difference was only significant within the peg-in-hole task. Table 6 also lists the average error score for each task in experiment 1.

Table 6. Averages and significant t-tests for machine differences in master/slave mode, experiment 1

Task	Average log time[a]			Average errors		
	BILARM	M-2	t-test	BILARM	M-2	t-test
All	1.651 (44.8)	1.326 (21.2)	18.70[b]	2.283	1.535	
Impact wrench	1.575 (37.6)	1.319 (20.8)	6.25[c]	2.260	1.940	
Peg in hole	1.927 (84.5)	1.447 (28.0)	20.21[b]	3.210	1.210	3.78[c]
Motor mount	1.457 (28.6)	1.209 (16.2)	10.21[b]	1.380	1.450	

[a]Numbers in parentheses are task times in seconds.
[b]Significant at $\alpha \leqslant 0.001$.
[c]Significant at $\alpha \leqslant 0.05$.

4.4.2 Force-Reflection Effects

There were no significant differences between force-reflection and nonforce-reflection conditions for time to completion or for the number of errors committed per trial for the M-2 or the BILARM in experiment 1.

4.4.3 Control Mode Differences Within the BILARM

A separate analysis compared the BILARM with master/slave control to the BILARM with switchbox control. For each task completion time, the BILARM was significantly faster in the master/slave mode than in the switchbox mode. Surprisingly, operators using the master/slave mode on the peg-in-hole task committed significantly more errors than they did using the switchbox mode, although there were no differences on the other tasks. Table 7 lists the averages of task time, the logarithm to base 10 of task time, and the errors on each task for both BILARM control system.

Table 7. Averages and significant t-tests for BILARM, experiment 1

Task	Average log time[a]			Average errors		
	Switch	M/S	t-test	Switch	M/S	t-test
All	2.046 (111.2)	1.651 (44.8)	7.99[b]	2.280	2.283	
Impact wrench	2.068 (117.0)	1.575 (37.6)	9.12[b]	3.260	2.260	
Peg in hole	2.194 (156.3)	1.927 (84.5)	4.78[b]	1.740	3.210	3.99[b]
Motor mount	1.875 (75.0)	1.457 (28.6)	6.85[b]	1.840	1.450	

[a]Numbers in parentheses are task times in seconds.
[b]Significant at $\alpha \leqslant 0.05$.

4.4.4 Machine Differences Within Switchbox Control System

When averaged across all three tasks, operators using the BILARM completed tasks in significantly less time than they did when using the PaR arm. However, there were no significant differences between these machines within individual tasks. There was a significant difference in the number of errors between these machines for all tasks combined and on the peg-in-hole task, with fewer errors committed by operators using the PaR arm. Table 8 lists averages and significant t-tests for time to completion, logarithm to base 10 of time to completion, and errors for the switchbox-controlled systems.

Table 8. Averages and significant t-tests for switchbox-controlled systems, experiment 1

Task	Average log time[a]			Average errors		
	BILARM	PaR	t-test	BILARM	PaR	t-test
All	2.046 (111.2)	2.086 (121.9)	6.65[b]	2.360	1.840	3.31[b]
Impact wrench	2.068 (117.0)	2.104 (127.0)		3.300	3.000	
Peg in hole	2.194 (156.3)	2.225 (167.9)		1.780	1.030	4.54[b]
Motor mount	1.875 (75.0)	1.929 (85.0)		2.000	1.480	

[a]Numbers in parentheses are task times in seconds.
[b]Significant at $\alpha \leqslant 0.05$.

4.4.5 Subtask Analysis

Time to completion of subtasks was analyzed by converting time in seconds to the percent of total time per subtask (Table 5 lists the subtasks). The purpose of this conversion was to allow direct comparison of the allocation of time to each subtask, independent of the actual time spent completing each task. Figure 13 shows the proportions of time allocated to each of the subtasks of the impact-wrench task for the six combinations of the manipulator system and control mode. The figure suggests that the operators allocated a larger proportion of their time to the third subtask (moving the wrench to the location of the bolt) when using the M-2 than when using other systems. However, the M-2 was generally faster; consequently, this difference could reflect the relative speed with which operators completed the other four subtasks. Table 8 also shows that when the operators were using the PaR 6000 manipulator, they allocated a relatively larger proportion of their time to the second subtask (alignment of the tongs and gripping the wrench). Subtask times were proportionally longer for the BILARM in the master/slave mode on the first and sixth subtasks (move to wrench and insert wrench) than for other machine and mode combinations.

Figure 14 shows the proportions of time allocated to each of the subtasks of the peg-in-hole task. This graph reveals several differences: (1) compared with the master/slave modes, the BILARM in switchbox mode required larger proportions of time for alignment of the tongs and gripping (subtask 2); (2) the BILARM in the master/slave mode required a relatively small allocation of time for extracting the peg (subtask 3); (3) the master/slave combinations all had relatively large allocations of time for alignment of the peg with the hold for reinsertion (subtask 5); and (4) the PaR arm required a larger proportion of time for moving to and from the touch panel (subtask 4).

Figure 15 shows the allocation of time in the motor-mount task. The BILARM and the PaR 6000 had relatively large proportions of time allocated to alignment and gripping, but relatively small allocations of time to dismounting the motor.

The results of the subtask analysis reveal an interesting pattern. The switchbox systems required relatively longer proportions of total task time to complete subtasks that required alignment of the tong with an object to be grasped. The numerous small adjustments in position necessary to align an end effector with an object may be indicative of the relative complexity (to the operator) and inefficiency of joint-by-joint control. This type of control requires a better perception (by operators) of the position of each manipulator link and of how the motion of each link will affect every other link and joint. This is a more complex task than "flying" an end effector in space, as is practiced with master/slave control. Joint-by-joint control is less efficient because it requires serial control of each joint and it is difficult for operators to control more than one joint at a time. Flying the end effector allows simultaneous control of every joint because the links and joints follow the motion of the operator's hand.

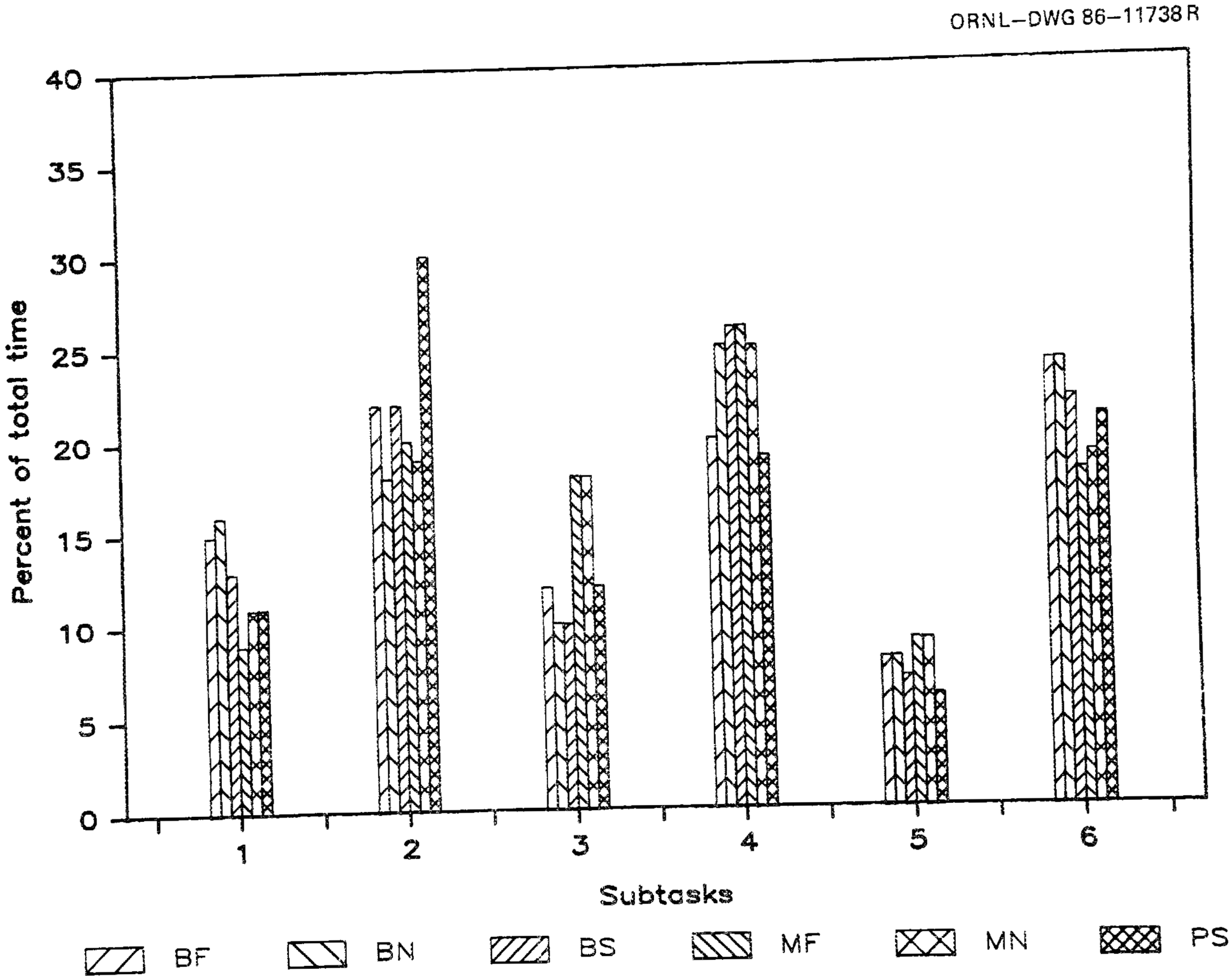

Fig. 13. Impact wrench task.

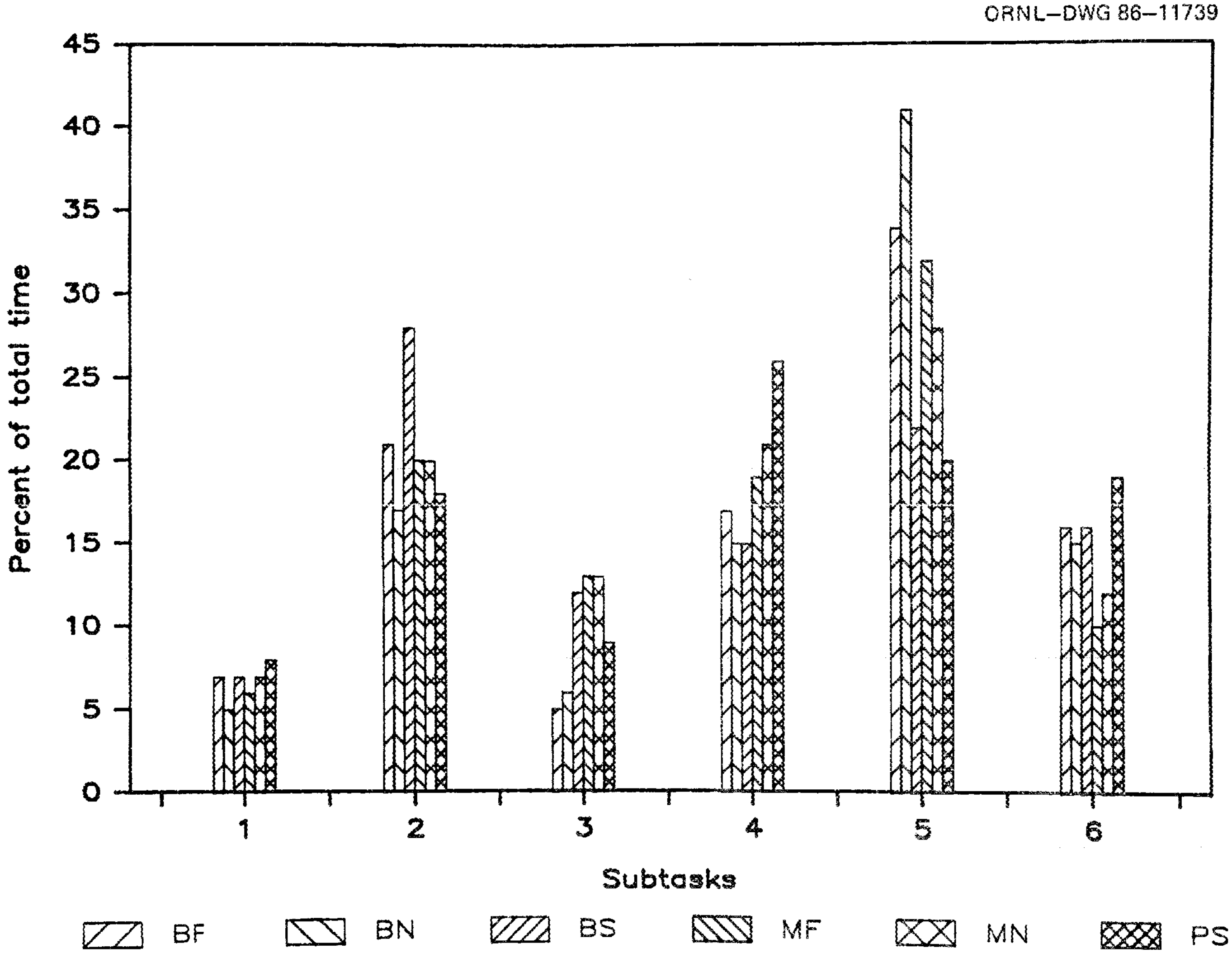

Fig. 14. Peg-in-hole task.

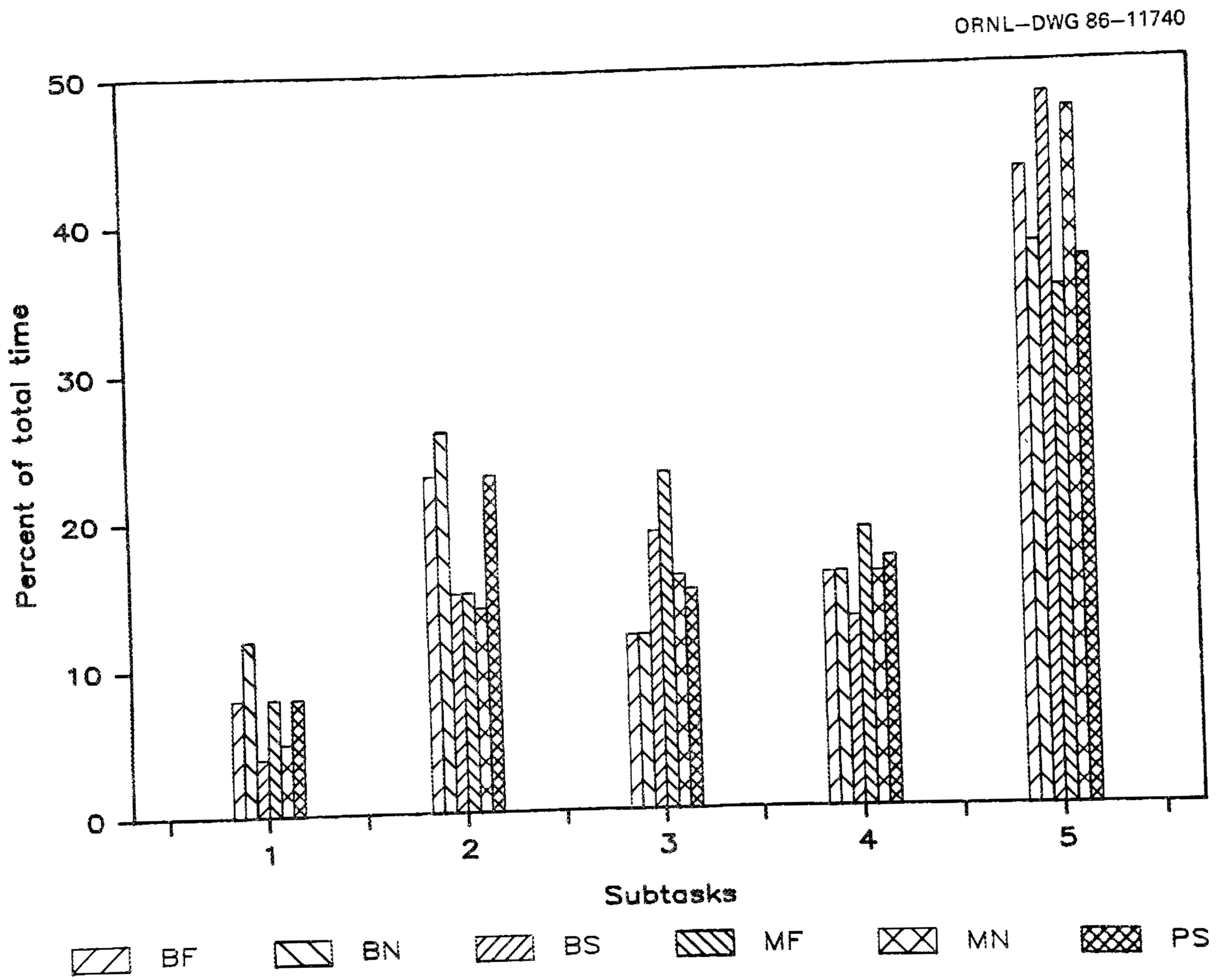

Fig. 15. Motor-mount task.

4.5 PRACTICE EFFECT

The five operators showed a consistent tendency to improve their performance over the course of testing. The average time to completion for the three tasks combined showed a fairly steady decline from the first to the tenth session. Only the ten sessions involving master/slave control were included; this was to avoid confounding the effect of practice with the difference between master/slave control and switchbox control. The linear trend (of the logarithm to base 10 of task time in seconds) was tested for statistical significance using a simple contrast, and the result was significant. For the linear trend across ten sessions on the average of the logarithm to base 10 of time to completion for the three tasks combined, $t = 3.09$ and $\alpha \leqslant 0.05$, which is significant. The trend toward faster performance with practice reflected a general tendency for operators to do better during their third or fourth session with a particular combination of manipulator and mode than during their first or second session with the same combination. If all operators had performed combinations in the same order, differences due to manipulator systems and modes would have been impossible to separate from differences due to practice. Figure 16 illustrates the downward trend in task completion time across sessions.

ORNL–DWG 86–11741

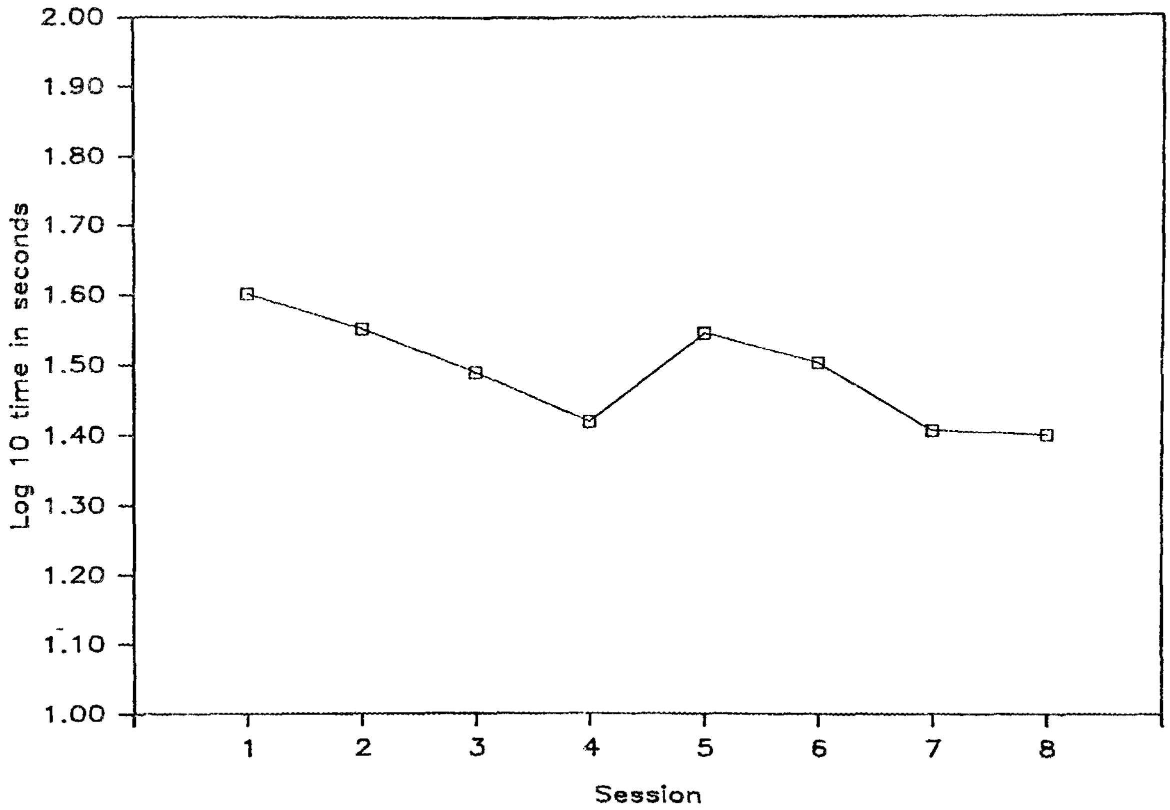

Fig. 16. Practice on master/slave.

5. EXPERIMENT 2

Operators using the M-2 completed tasks in shorter times and committed fewer critical incidents than they did with the BILARM. The effect of force reflection was mainly confined to the BILARM; with force reflection, time to completion was usually greater, and there were generally more critical incidents. There was no effect of force reflection with the M-2. The BILARM completed tasks in less time with master/slave control than with switchbox control. The PaR arm was faster than the BILARM on the peg-in-hole task.

5.1 TASK DESCRIPTIONS

The three tasks selected for experiment 2 (Fig. 17) were designed to be representative of remote handling task primitives for recycle facilities and other general remote handling

ORNL–PHOTO 2485–85

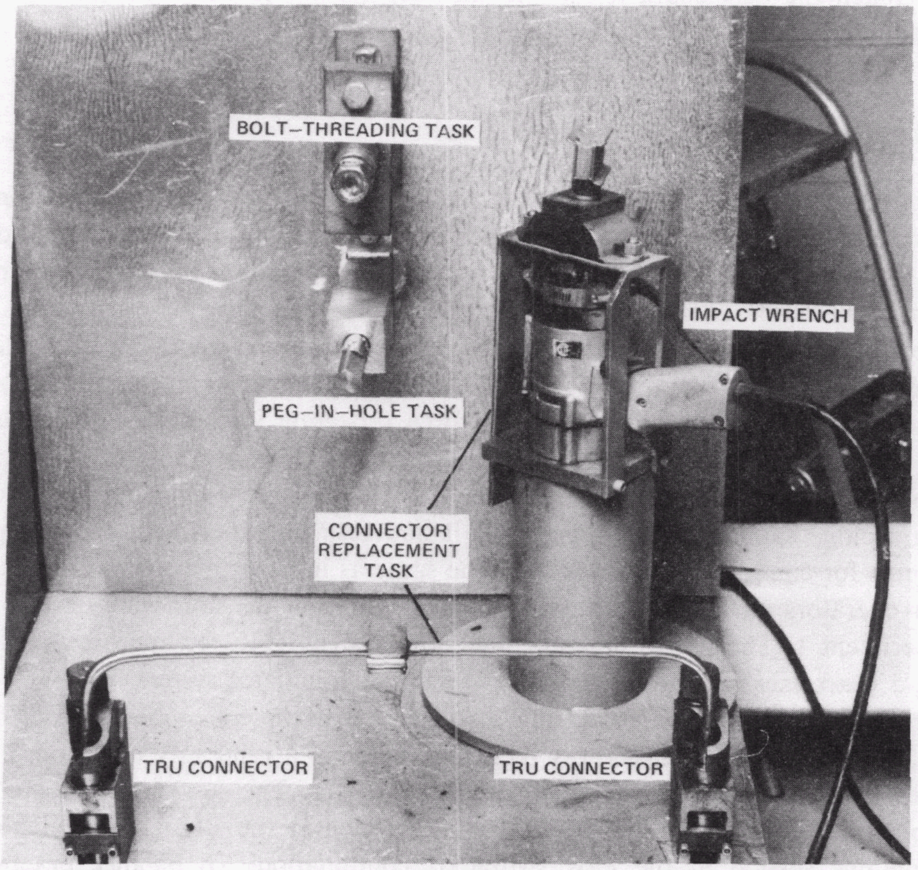

Fig. 17. Tasks used in experiment 2.

applications. They were also designed to be easily completed with each of the combinations of manipulator and control mode. Each task was simple enough to be completed several times in a 1-h testing session; however, they were somewhat more difficult than the tasks in experiment 1. The following sections describe the tasks in detail.

5.1.1 Bolt Threading

In this task the operator used the manipulator to grasp a hexagonal-headed, captive bolt that protruded horizontally from the vertical task board. The threads of the bolt were engaged with the threads in a socket, and the bolt was rotated until fully inserted. The bolt had a machined lead-in at the end to ease insertion. When the bolt reached the criterion depth for task completion, it closed a microswitch and actuated a signal light. The trial was complete when the signal light came on.

5.1.2 Connector Replacement

In this task operators used the manipulator arm to grasp a power impact wrench and loosen two TRU connectors that held the ends of a tubing jumper in their sockets. After the TRU connectors were loosened, the tubing jumper was removed from its sockets, moved to touch the task board shelf, and then replaced. The trial was complete when the ends of the jumper were firmly in their sockets.

5.1.3 Peg in Hole (modified)

During experiment 1, the panel that the peg-in-hole task was mounted on was not rigidly anchored to the task board; thus, it flexed as the peg was extracted and inserted. The flexing allowed the task to be completed with less precision than would be required if the mount were rigid. The motion of the panel provided a form of visual force feedback. The degree of displacement of the panel provided information about forces applied by the manipulators. It is possible that this unwanted force feedback in the nonforce-reflection condition artificially reduced the effects of force reflection in the peg-in-hole task. Therefore, the mounting panel was stiffened in experiment 2 to prevent flexing under forces applied by the manipulator systems. The peg had two flat faces cut onto opposite sides for ease of grasping, but it was not tapered as it had been in experiment 1. Procedures and requirements for completing this task were the same as in experiment 1.

Four operators participated in experiment 2. Three of the four were part of the group from experiment 1. The fourth operator had worked at ORNL for approximately 38 years, including 5 years as a part-time operator of remote manipulator systems.

Each operator completed 18 testing sessions, and each session consisted of 6 trials. In a session, each of the three tasks was performed once in a randomly determined order and then repeated in a different randomly determined order. The session series was organized as four blocks of four sessions, followed by one session in the BILARM switchbox combination and one session in the PaR switchbox combination. To balance practice effects, within each block, every operator completed all four of the master/slave combinations (BILARM force reflecting, BILARM nonforce reflecting, M-2 force reflecting, M-2 non-force reflecting). In each of the four blocks, the four master/slave combinations were arranged in a different order.

The testing sequence was specially designed to equalize the effects of transferring from one combination into another. Each operator's sequence of 16 sessions of master/slave combinations included all of the possible types of changes from one condition to another. For instance, the change from the M-2 with force reflection to the BILARM with force reflection was one of 12 possible changes. The opposite change (from BILARM with force reflection to the M-2 with force reflection) was another. Table 9 shows the testing sequences for the four operators who participated in experiment 2.

Table 9. Testing sequence in experiment 2

Operator	Sequence									Sequence								
1	BN[a]	MF[b]	MN[c]	BF[d]	MN	MF	BF	BN	BF	MF	BN	MN	BF	MF	MN	BN	BS[e]	PS[f]
2	BF	MF	MN	BN	MF	BF	MN	BN	BF	BN	MN	MF	BN	MF	MN	BF	BS	PS
3	MN	BF	BN	MF	BN	BF	MF	MN	MF	BF	MN	BN	MF	BF	BN	MN	BS	PS
4	MF	BF	BN	MN	BF	MF	BN	MN	MF	MN	BN	BF	MN	BF	BN	MF	BS	PS

[a]BILARM nonforce reflecting.
[b]M-2 force reflecting.
[c]M-2 nonforce reflecting.
[d]BILARM force reflecting.
[e]BILARM with switchbox.
[f]PaR manipulator.

5.2 PROCEDURES

Before beginning the testing trials in experiment 2, the four operators were given 1 h of practice on the M-2 manipulator and 3 h on the BILARM with master/slave controls. They were also given 1 h of practice with each of the two switchbox manipulators.

Prior to testing, the operators were also given a set of instructions that emphasized the importance of working for precision as well as for speed. The instructions were as follows:

> In this test you will perform three tasks with each manipulator system. The purpose of these tasks is to compare the performance of the manipulators in as realistic a setting as possible; that is, under conditions similar to those normally required in an operational fuel recycling plant.
>
> In this test, performance will be measured in several ways: (1) the number and type of errors you make while performing the tasks will be recorded, (2) you will fill out an evaluation of how difficult it is to perform the tasks with each manipulator and feedback condition (that is, force or nonforce), and (3) the time required to perform each task will also be recorded. However, do not rush through the tasks as if you were in a race. It is important that you attempt to avoid making errors that could lead to equipment damage in a real facility. Work as quickly as you can but do not sacrifice the quality of your work for speed.

The instructions also included the definitions of errors that were used in experiment 1: misalignments, misses, drops, and overloads/damage. The list of critical incidents was not given to the operators at this time to avoid contaminating their responses in interviews conducted at a later date. Instructions were read aloud and presented in writing. At the end of each session, operators were asked to complete a questionnaire in which they rated the difficulty they experienced in completing the session. The questionnaire is reproduced in Table 10. Each session was followed by a rest period of at least 1 h.

The procedure followed during testing sessions in experiment 2 was the same as the procedure followed in experiment 1.

Table 10. Task difficulty questionnaire

1.______	Very easy; required minimal levels of effort
2.______	Easy; required low levels of effort
3.______	Mildly difficult; required acceptable levels of effort
4.______	Difficult; required moderately high effort
5.______	Moderately difficult; required high levels of effort
6.______	Very difficult; required very high effort
7.______	Extremely difficult; required intense effort

5.3 DEPENDENT VARIABLES

Time to completion was defined in the same way for experiment 2 as it was for experiment 1. Errors were not used as a dependent variable in experiment 2; instead of the testing supervisor recording the occurrence of the limited set of errors as in experiment 1, an observer viewed each of the videotapes of the testing sessions in experiment 2. The observer recorded the occurrence of 18 different critical incidents during task performance. Critical incidents are events or behaviors that reflect the quality of performance with a manipulator system. An analysis of individual critical incidents revealed that four of the incidents did not occur frequently enough to be used; these were stopping task (17), damage (13 and 14), and tangling (11). The critical incidents used in experiment 2 are described in the following sections.

5.3.1 Miscentering (1)

This category was defined as the incorrect location of the point of contact between the tong of the manipulator and a component of the task, or between a grasped object and a component. For example, when an operator moved the peg toward the hole and the peg made contact with the surface near the hole, but not squarely in it, the incident was counted as miscentering.

5.3.2 Misalignment (2 and 3)

This incident occurred when the axis of motion of an object held by the manipulator was not aligned as required by the task. For example, when an operator attempted to insert the peg into the hole, but the axis of peg motion was not the same as the axis of the hole, an incident of misalignment was scored. Misalignments were scored separately when they occurred during insertion of an object into a receptacle (misalign-insertion, 2) or during extraction from a receptacle (misalign-extraction, 3).

5.3.3 Slipping (4 and 5)

Slipping was cited when an item held by the tong slid out of position but was not completely released. For example, when the handle of the impact wrench slid from the back of the tong to the front, the incident was scored as a slip (if the wrench actually left the tong, the incident was scored as dropping). Slipping was scored separately when the operator continued without repositioning the item (slip-continue, 4) or stopped to reposition the item (slip-reposition, 5).

5.3.4 Using Opposite Hand (6)

This incident occurred when an operator used a left hand (all of the operators were right handed) to support or guide the master arm or to hold the tongs closed.

5.3.5 Collision (7 and 8)

This incident was defined as accidental contact between any part of the manipulator (or an object in its grasp) and any object in the task area. For example, when the manipulator arm bumped into the task board an incident of collision was scored. Collisions specifically excluded intentional contacts, such as contact between the impact wrench and a bolt during the connector task. Collisions were scored separately when the contact involved the manipulator itself (collision-manipulator 7) or a grasped object (collision-grasped object, 8).

5.3.6 Missing (9)

This incident occurred when an attempt to make contact with an object failed and no contact was established, for example, when an operator attempted to grasp the connector and closed the tongs without touching it.

5.3.7 Pressing (10)

When an operator pushed or pulled some portion of the task area with enough force to cause visible deflection of the mounting board or any part of the task, a pressing incident was recorded.

5.3.8 Tangling (11)

This incident was counted when the slave portion of the manipulator became entangled with or caught on some object in the task area. For example, when the tong of the manipulator arm was unintentionally hooked behind the connector tube, the incident was scored as tangling.

5.3.9 Groping (12)

Any occurrence of a repeated series of attempts to correctly position the tong or an item in the grasp of the tong (such as the impact wrench) was defined as groping. For example, groping was scored when an operator made a series of attempts to position the tong so that he could grasp the bolt in the bolt-threading task.

5.3.10 Damage (13 and 14)

Whenever the manipulator or an object in the testing area was bent, scratched, gouged, broken, or otherwise damaged during the performance of a test trial, a damage incident was recorded. Damage was scored separately for damage to the manipulator (damage-manipulator, 13) and to other objects (damage-object, 14).

5.3.11 Dropping (15)

This incident was cited when an object held by the tongs slipped out of the control of the tong, the exception being when the task called for the release of an object. For example, when the impact wrench dropped from the tong while it was being moved to the task board, an incident of dropping was scored.

5.3.12 Incomplete Release (16)

This incident occurred when an operator released a grasped object as required by the task, but the tong remained in contact with the object. For example, when an operator attempted to release the bolt at the appropriate point in the bolt-threading task, but one claw of the tongs stayed in contact with the bolt, an incident of incomplete release was scored.

5.3.13 Stopping the Task (17)

This event included times when an incident occurred that required ending a trial before the task was completed. For example, if the manipulator arm was damaged and had to be repaired before the trial could continue, an incident of stopping would have been scored.

5.3.14 Other Incident (18)

This category was defined as any incident that the observer judged important to performance of the task but for which there was no predefined category. The observer recorded the nature of the incident whenever a similar incident occurred.

Each of the three tasks was analytically divided into an exhaustive series of sequential subtasks. One of the two completions per session of each task (randomly selected) was analyzed by subtask. The sample included three of the six trials in each session. From the videotapes made of each session, an observer recorded the amount of time taken to complete each subtask for each of the trials selected. Table 11 lists the subtasks for each task

Table 11. Subtasks for the three tasks used in experiment 2

	Task		
Subtask	Bolt threading	Peg in hole	Connector
1	Move from start to bolt turning	Move from start to peg holder	Move from start to wrench, grasp and remove
2	Bolt turning to final release	Align tong, grasp peg, and extract	Move to bolt 1 and loosen
3		Move to panel and touch	Loosen bolt 2
4		Move to peg holder, align, and insert	Replace wrench
5			Grasp jumper and remove from mount
6			Move to touch panel
7			Replace jumper

in experiment 2. For the peg-in-hole task, two of the subtasks defined in experiment 1 were combined with other subtasks to simplify the analysis.

5.4 RESULTS

This section discusses the results associated with each dependent variable for comparisons among the six combinations of manipulator system and mode of control. Results for the master/slave combinations are presented first, followed by switchbox combinations and a comparison between BILARM master/slave and switchbox modes. Overall results of the subtask analysis follow, concluding with results of the test for practice effects.

Figure 18 shows the averages of task completion time, and Fig. 19 shows the total critical incidents for the machine and mode combinations in experiment 2. Figure 20 shows the average difficulty rating for each machine and mode combination. The average difficulty rating was the average of each operator's responses to the questionnaire. A high score indicates that the operators perceived the control and mode combination as difficult to use, and a low score indicates that the operators felt it was easy to use.

ORNL–DWG 86–11742

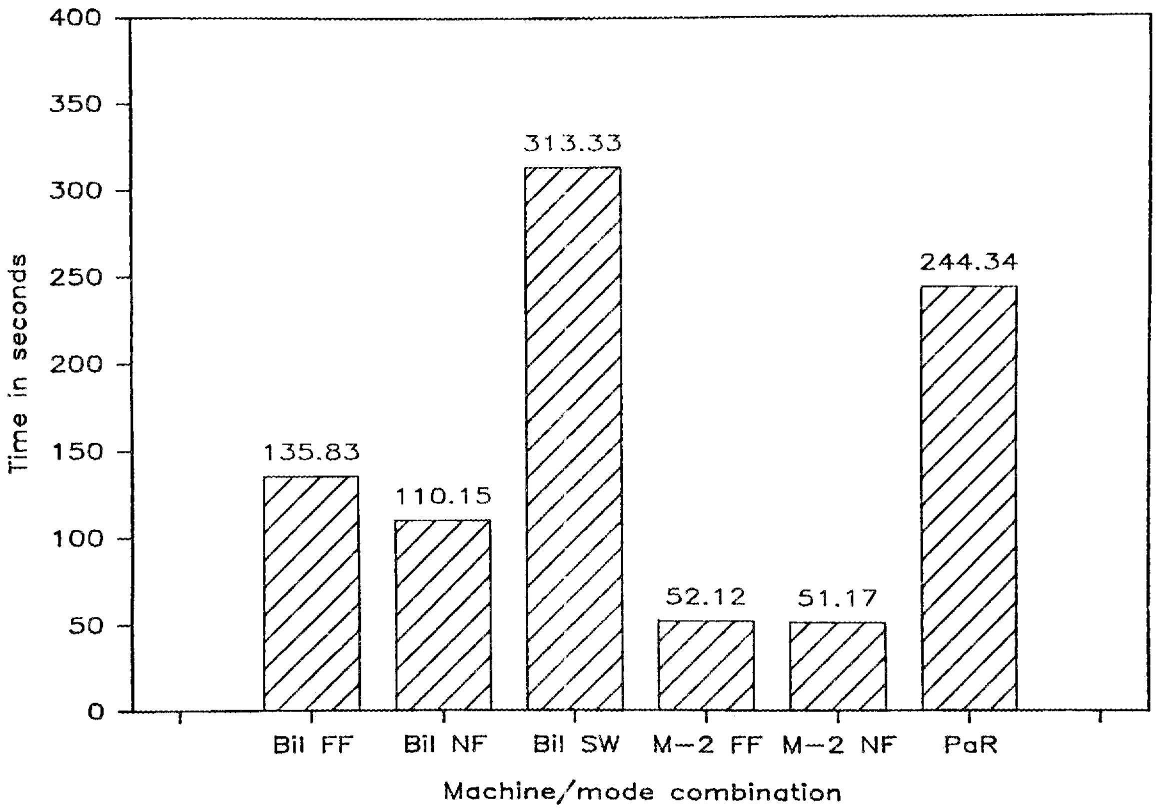

Fig. 18. Time to complete tasks.

ORNL–DWG 86–11743

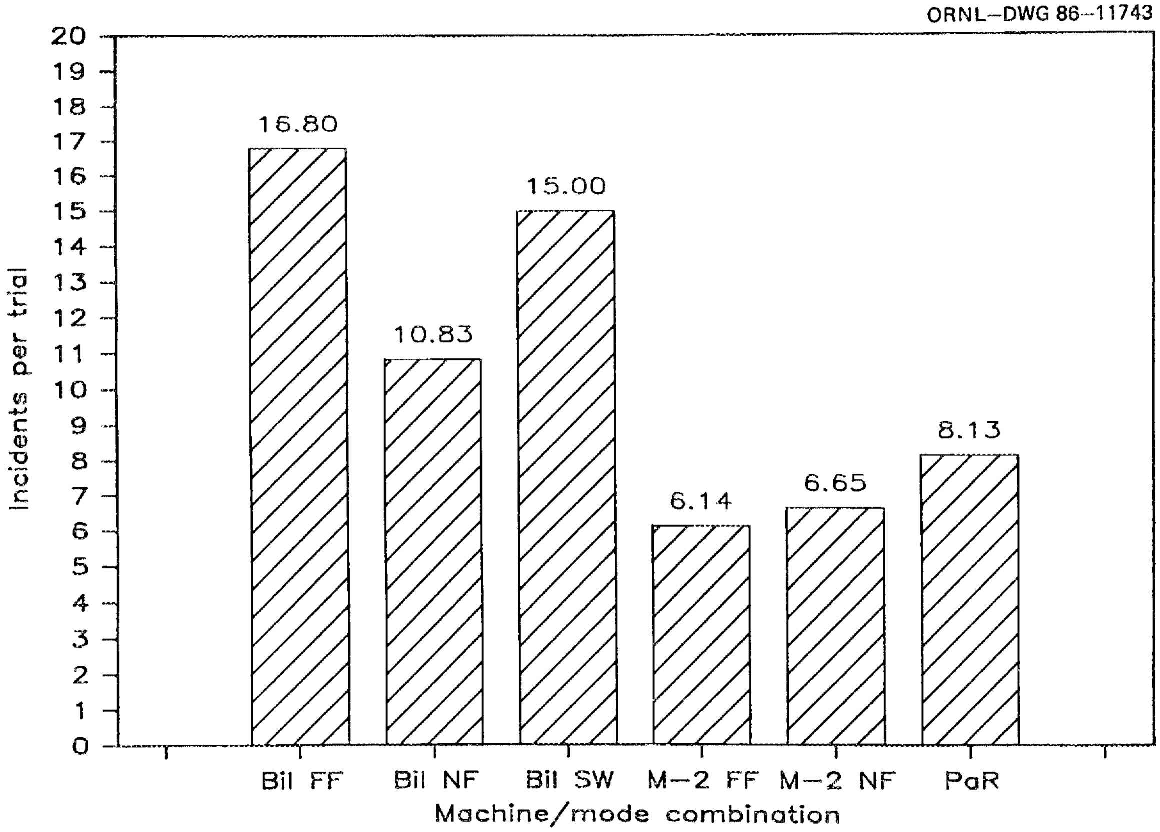

Fig. 19. Total incidents per trial.

ORNL–DWG 86–11744

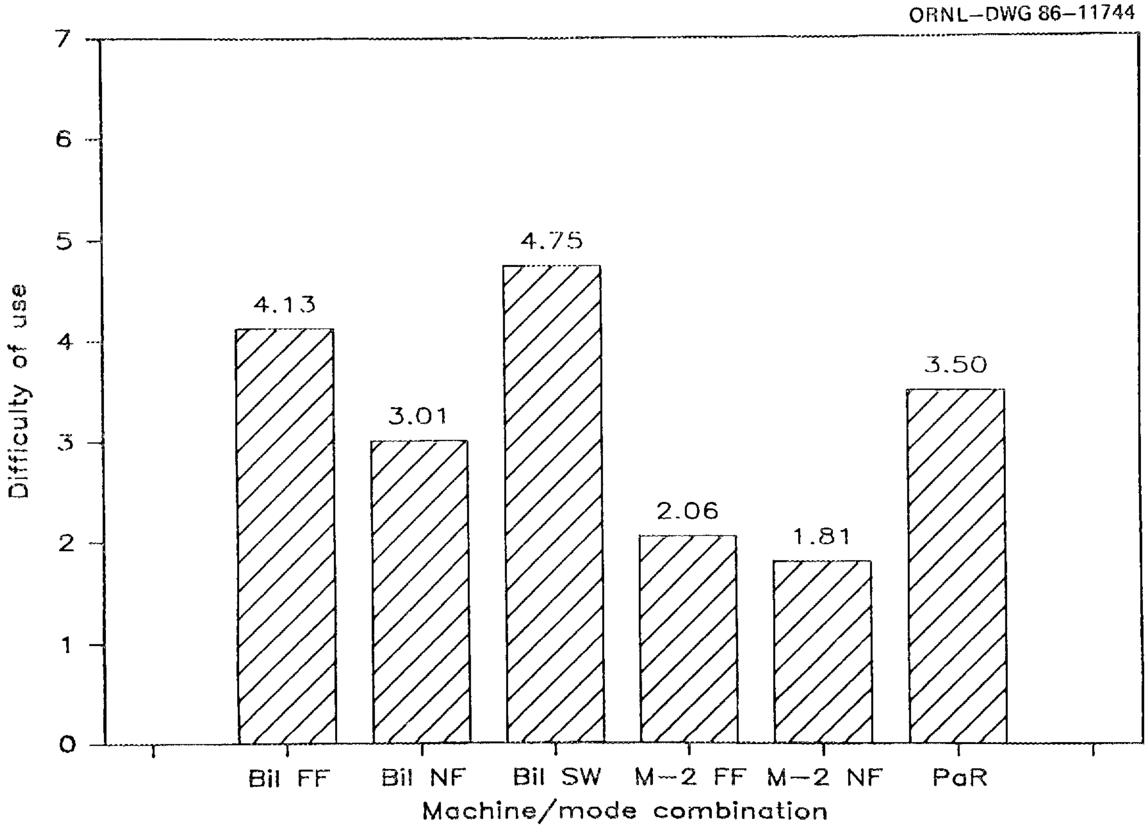

Fig. 20. Ratings of difficulty.

5.4.1 Machine Differences Within Master/Slave Systems

Operators using the BILARM in all three tasks required significantly more time to complete tasks than did operators using the M-2, as shown in Table 12. The number of critical incidents was significantly greater for the BILARM than for the M-2, that is, for the average of all three tasks and within each task. For all tasks, critical incidents that revealed significant differences between machines occurred more frequently on the BILARM than on the M-2. Within each task, certain incidents did not occur frequently enough to be used in the analysis of individual incidents. Table 13 through 15 list incidents that showed significant differences in each task, along with t-tests comparing the BILARM in master slave mode with the M-2 for each incident. Operators rated the BILARM significantly higher than the M-2 on the task difficulty questionnaire ($t = 3.88$, $\alpha \leqslant 0.05$).

Table 12. Averages and significant t-tests for machine differences in master/slave mode, experiment 2

Task	Average log time[a]		
	BILARM	M-2	t-test
All	2.088 (122.5)	1.715 (51.9)	8.92[b]
Bolt threading	2.134 (136.1)	1.555 (35.9)	20.18[c]
Peg in hole	1.939 (86.9)	1.596 (38.4)	4.13[b]
Jumper	2.189 (154.5)	1.939 (86.9)	5.70[b]

[a]Numbers in parentheses are task times in seconds.
[b]Significant at $\alpha \leqslant 0.05$.
[c]Significant at $\alpha \leqslant 0.001$.

Table 13. Average critical incidents and significant t-tests in master/slave mode, experiment 2 (bolt-threading task)

Incident	BILARM	M-2	t-test
Total	19.38	5.30	5.50[a]
Pressing	4.17	3.15	5.58[a]
Dropping	9.13	1.85	5.09[a]

[a]Significant at $\alpha \leqslant 0.05$.

Table 14. Average critical incidents and significant t-tests in master/slave mode, experiment 2 (peg-in-hole task)

Incident	BILARM	M-2	t-test
Total	8.36	4.97	4.08[a]
Miscenter	1.06	0.63	5.17[a]
Slip-reposition	0.74	0.22	9.66[b]
Collision manipulator	0.31	0.03	6.97[b]
Press	1.25	0.12	5.23[a]
Dropping	1.53	0.70	4.14[a]

[a]Significant at $\alpha \leqslant 0.05$.
[b]Significant at $\alpha \leqslant 0.01$.

Table 15. Average critical incidents and significant t-tests in master/slave mode, experiment 2 (connector-replacement task)

Incident	BILARM	M-2	t-test
Total	13.22	8.92	6.69[a]
Misalign-insert	1.92	1.06	5.06[b]
Opposite hand	0.89	0.78	7.00[a]
Miss	1.11	0.44	3.51[b]
Dropping	1.53	0.70	4.14[b]
Incomplete release	0.36	0.06	19.00[a]

[a]Significant at $\alpha \leqslant 0.05$.
[b]Significant at $\alpha \leqslant 0.01$.

5.4.2 Force-Reflection Effect

In general, BILARM performance was better under the nonforce-reflection condition. In contrast, for the M-2, there was no consistent difference in operating performance between force reflection and nonforce reflection. The time to completion and the logarithm to base 10 of time to completion for each task for each machine are shown in Table 16.

The effect of force reflection on task completion time was not significant in the connector-replacement task; however, there was a significant difference in the way the machines reacted to force reflection ($t = 4.22$, D.F. $= 3$, $\alpha \leqslant 0.05$). Figure 21 shows that differences due to force reflection occurred almost entirely on the BILARM.

For the BILARM, the number of critical incidents was consistently greater in the force-reflection condition. The effect of force reflection with the M-2 was inconsistent. In some cases force reflection led to fewer errors; in other cases it led to more errors.

In the bolt-threading task, the effect of force reflection was detrimental with the BILARM, but there was no effect with the M-2. In particular, there were significantly

Table 16. Averages for logarithm to base 10 of time to completion for force-reflection differences in master/slave mode, experiment 2[a]

Task	BILARM		M-2	
	Force	Nonforce	Force	Nonforce
All	2.133 (135.8)	2.042 (110.2)	1.717 (52.1)	1.714 (51.8)
Bolt threading	2.244 (175.4)	2.023 (105.4)	1.621 (41.8)	1.589 (38.8)
Peg in hole	1.939 (86.9)	1.939 (86.9)	1.603 (40.1)	1.601 (39.9)
Connector replacement	2.215 (164.1)	2.164 (145.9)	1.928 (84.7)	1.951 (89.3)

[a]Numbers in parentheses are times in seconds.

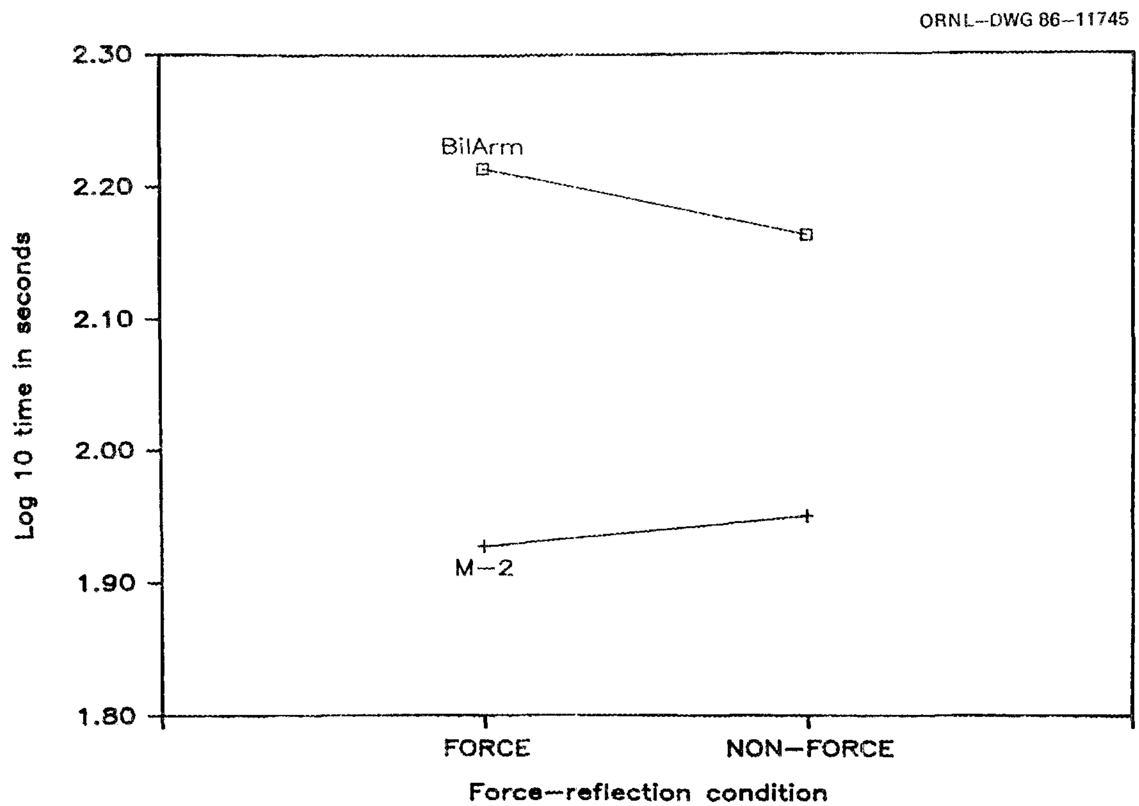

Fig. 21. Interaction, connector task.

more collision-manipulator and dropping incidents for force reflection with the BILARM. There was a significant difference between machines in the way that force-reflection conditions affected operators' performance (detrimental with the BILARM, but no significant difference with the M-2). This was also true for the total number of critical incidents ($t = 3.82$, $\alpha \leq 0.05$), collision-manipulator incidents ($t = 3.45$, $\alpha \leq 0.05$), and dropping incidents (t = 5.07, $\alpha \leqslant 0.05$); for other incidents, however, differences were not significant. Figures 22 through 24 illustrate the differences between machine and force-reflection combinations for these incidents. Although these illustrations show that force reflection led to fewer incidents per trial with the M-2, the differences were not statistically significant.

In the modified peg-in-hole task, force reflection led to more frequent total critical incidents, misalignment-insertion incidents, and slip-reposition incidents. Force reflection led to significantly fewer occurrences of the groping incident. Table 17 lists the average occurrence of significantly different critical incidents per trial in each force-reflection condition. As was the case for the other tasks, the detrimental effect of force reflection was confined to the BILARM; differences within the M-2 were slight. Significant differences in the way the machines reacted to force reflection were found for slip-reposition (t = 4.33, $\alpha \leqslant 0.05$) and slip-continue (t = 5.20, $\alpha \leqslant 0.05$). Figures 25 and 26 illustrate these differences. There seems to be an advantage for force reflection with the M-2 in terms of these incidents, but the difference between force-reflection and nonforce-reflection averages is not statistically significant.

In the connector-replacement task, total critical incidents were significantly higher with force reflection (t = 3.89, $\alpha \leqslant 0.05$); again, however, the increase in incidents was due primarily to the effect of force reflection on performance with the BILARM. Machine reaction to force-reflection conditions was significantly different for the "missing" incident. For this critical incident, there was a negative effect of force reflection with the BILARM but no effect with the M-2 (Fig. 27).

ORNL–DWG 86–11746

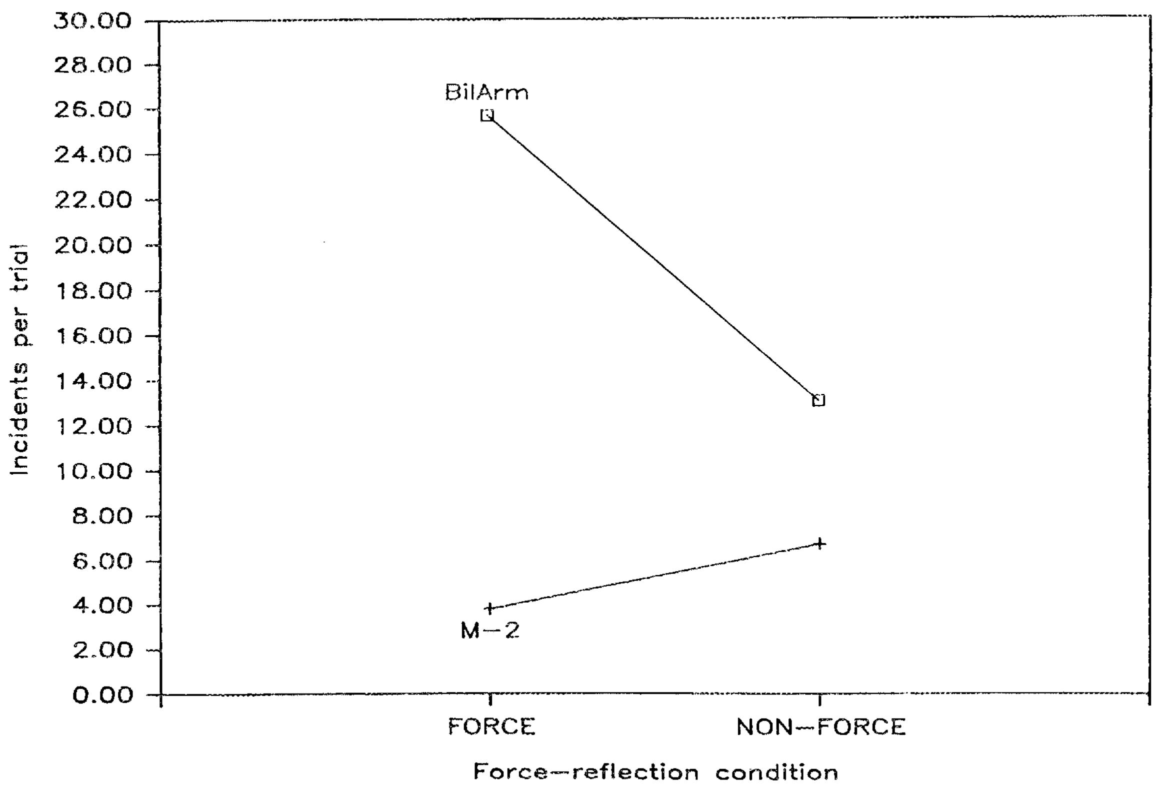

Fig. 22. Interaction, bolt-threading task.

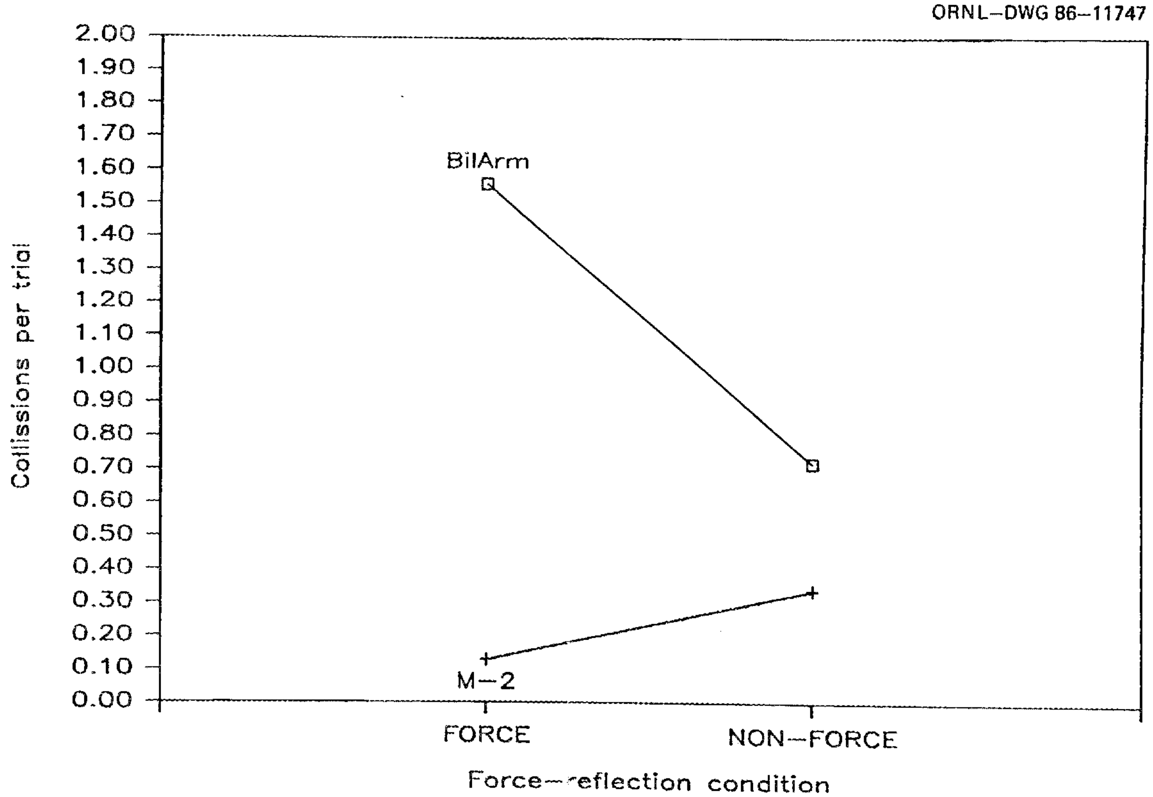

Fig. 23. Interaction, bolt-threading task.

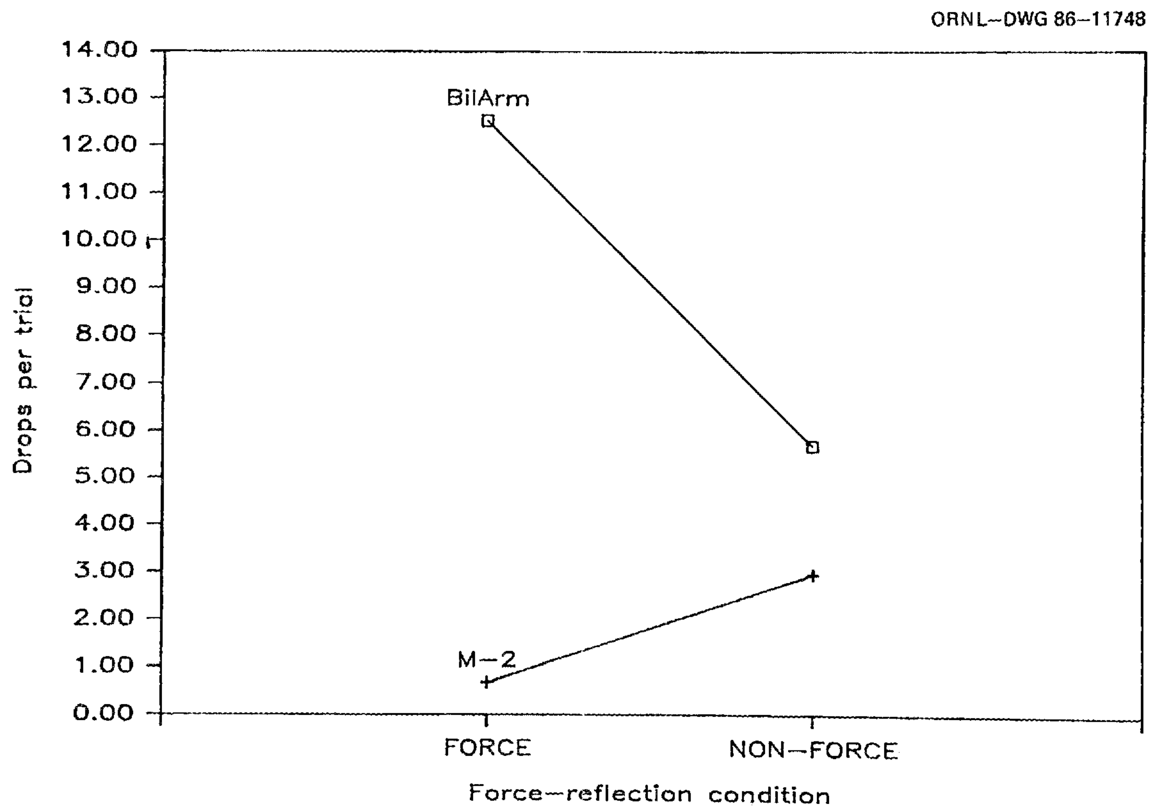

Fig. 24. Interaction, bolt-threading task.

Table 17. Average critical incidents and significant t-tests for force-reflection conditions, experiment 2 (peg-in-hole task)

Incident	Force reflection	Nonforce	t-test
Total	6.98	6.34	5.80[a]
Misalign-insert	1.04	0.64	3.81[a]
Slip-reposition	0.64	0.31	3.81[a]
Groping	1.02	1.16	3.58[a]

[a]Significant at $\alpha \leqslant 0.05$.

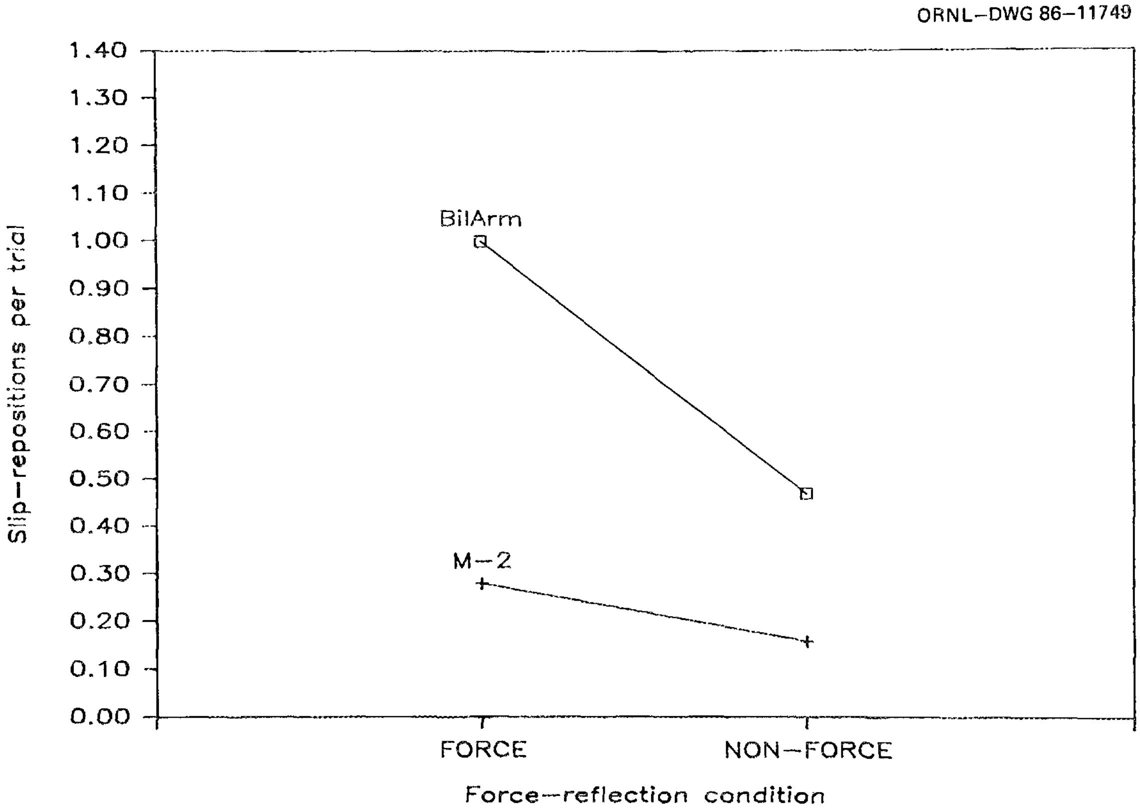

Fig. 25. Interaction, peg-in-hole task.

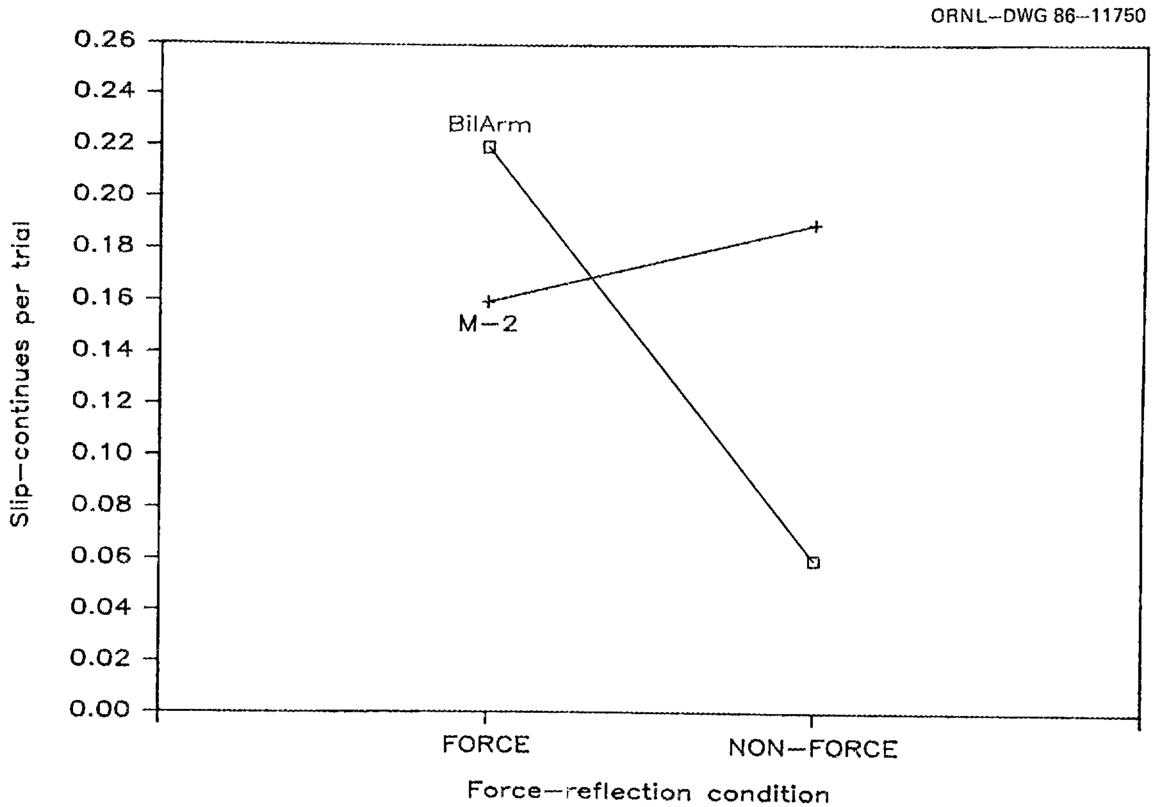

Fig. 26. Interaction, peg-in-hole task.

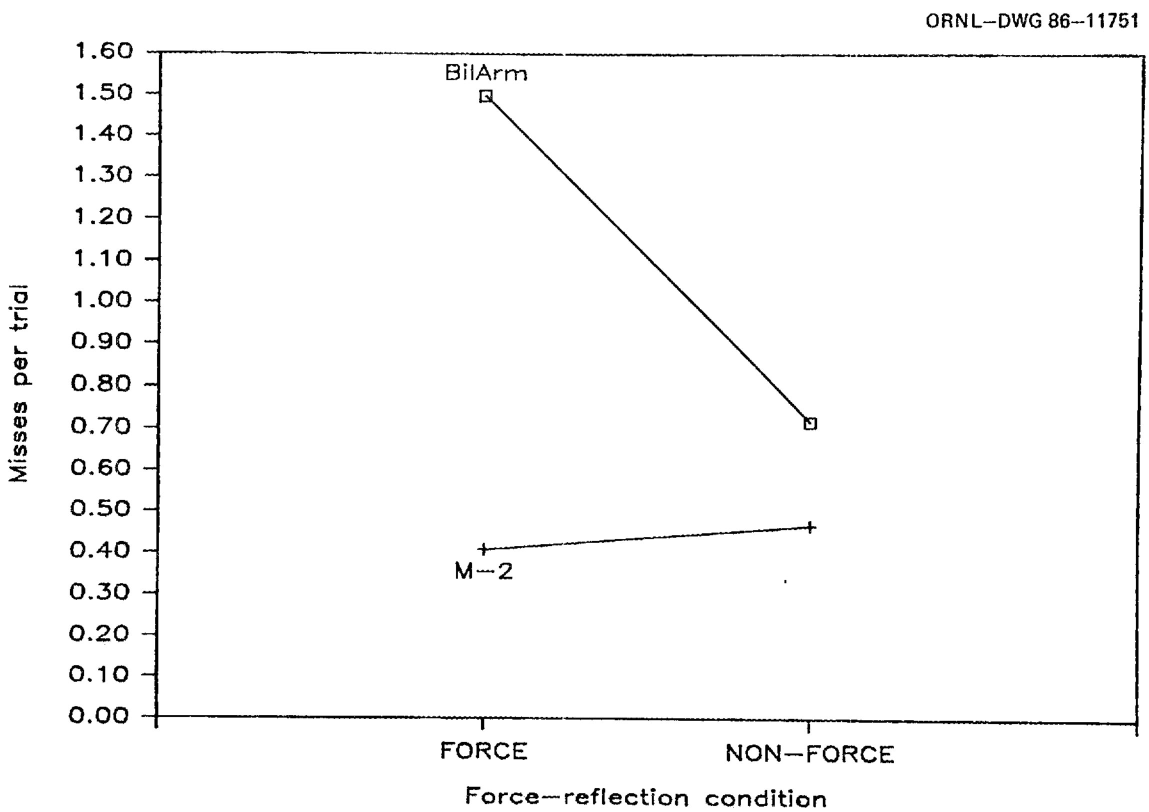

Fig. 27. Interaction, jumper task.

5.4.3 Control Mode Differences Within the BILARM

The analysis of control mode differences (master/slave vs switchbox control) within the BILARM concentrated on the task-completion-time variable. Critical incidents rarely occurred in trials involving switchbox-controlled manipulators; thus, they were considered too unreliable to be used in the analysis of these systems.

Operators using switchbox controls required more time to complete tasks than they did when using master/slave controls. Table 18 lists the averages of time to completion and t-test results.

Table 18. Averages and significant t-tests for control mode differences with the BILARM, experiment 2

Task	Average log time[a] M/S[b]	SWB[c]	t-test
All	2.088 (122.5)	2.496 (313.3)	8.47[d]
Bolt threading	2.134 (136.1)	2.439 (274.8)	5.92[e]
Peg in hole	1.939 (86.9)	2.398 (250.0)	5.27[e]
Jumper	2.189 (154.5)	2.649 (445.7)	10.34[d]

[a]Numbers in parentheses are times in seconds.
[b]Master/slave.
[c]Switchbox control.
[d]Significant at $\alpha \leqslant 0.05$.
[e]Significant at $\alpha \leqslant 0.001$.

5.4.4 Machine Differences Within Switchbox Control Systems

The analysis of differences between switchbox-controlled systems centered on task completion time. Critical incidents were not included. There was a significant difference between operators' averages with the PaR arm vs the BILARM (see Fig. 18). This was also true for the bolt-threading task ($t = 3.85$, $\alpha \leqslant 0.05$). In both cases the operators were faster with the PaR arm than with the BILARM. However, for the other tasks the difference was not significant, which suggests that the difference between the two machines is related to the way the machines reacted to a specific characteristic of the bolt-threading task. The bolt-threading task was easier to complete with the PaR arm because of its continuous wrist roll. With the PaR arm, operators were able to extend the wrist along a straight line and then rotate the wrist continuously. Rotation of the BILARM was limited to $\pm 115°$ (see Table 1); therefore, operators using the BILARM had to tighten the bolt with a series of ratcheting motions. On tasks that were not dependent on this one motion for performance, there was no difference between switchbox-controlled manipulators.

5.4.5 Subtask Analysis

Subtask analyses were conducted separately for each task. In addition to the bar charts used in the analysis of subtask data in experiment 1, analysis of variance was conducted for each subtask. Figures 28 through 30 show the differences in the proportion of task time allocated to each subtask in experiment 2.

In the bolt-threading task, there were no statistically significant difference between the machine and mode combinations in the allocation of time to subtasks.

In the peg-in-hole task, operators allocated a significantly greater proportion of time to performing the second subtask (first contact to removal of peg from holder) with the M-2 than with the BILARM (master/slave mode) ($t = 3.59$, $\alpha \leqslant 0.05$). However, the M-2 was generally faster than the BILARM; this difference could reflect the absolute speed with which the operators completed the other subtasks. There was no difference in time between the BILARM in master/slave mode and the BILARM in switchbox mode on any of the subtasks. Also, there was no difference in time between the BILARM in switchbox mode and the PaR manipulator. As Fig. 29 illustrates, there were large average time differences between these systems; however, these differences were not stable enough among operators to be declared statistically significant.

In the connector-replacement task, there was a significant difference ($t = 7.25$, $\alpha \leqslant 0.01$) between the BILARM in master slave and the M-2 on the third subtask (loosening bolt 1 and 2). Operators using the M-2 allocated a larger proportion of their time to completing this subtask. This probably reflects the difference in speed on the other subtasks; the M-2 was generally much faster. For the first subtask (start to the complete removal of the wrench from its holder), there was a significant effect of force reflection ($t = 7.39$, $\alpha \leqslant 0.01$), with relatively more time allocated to its performance. This was also true for the BILARM and the M-2. There was no statistically significant difference between the BILARM in master/slave mode and the BILARM in switchbox mode on any of the subtasks. Also, there was no significant difference between the BILARM in switchbox mode and the PaR manipulator.

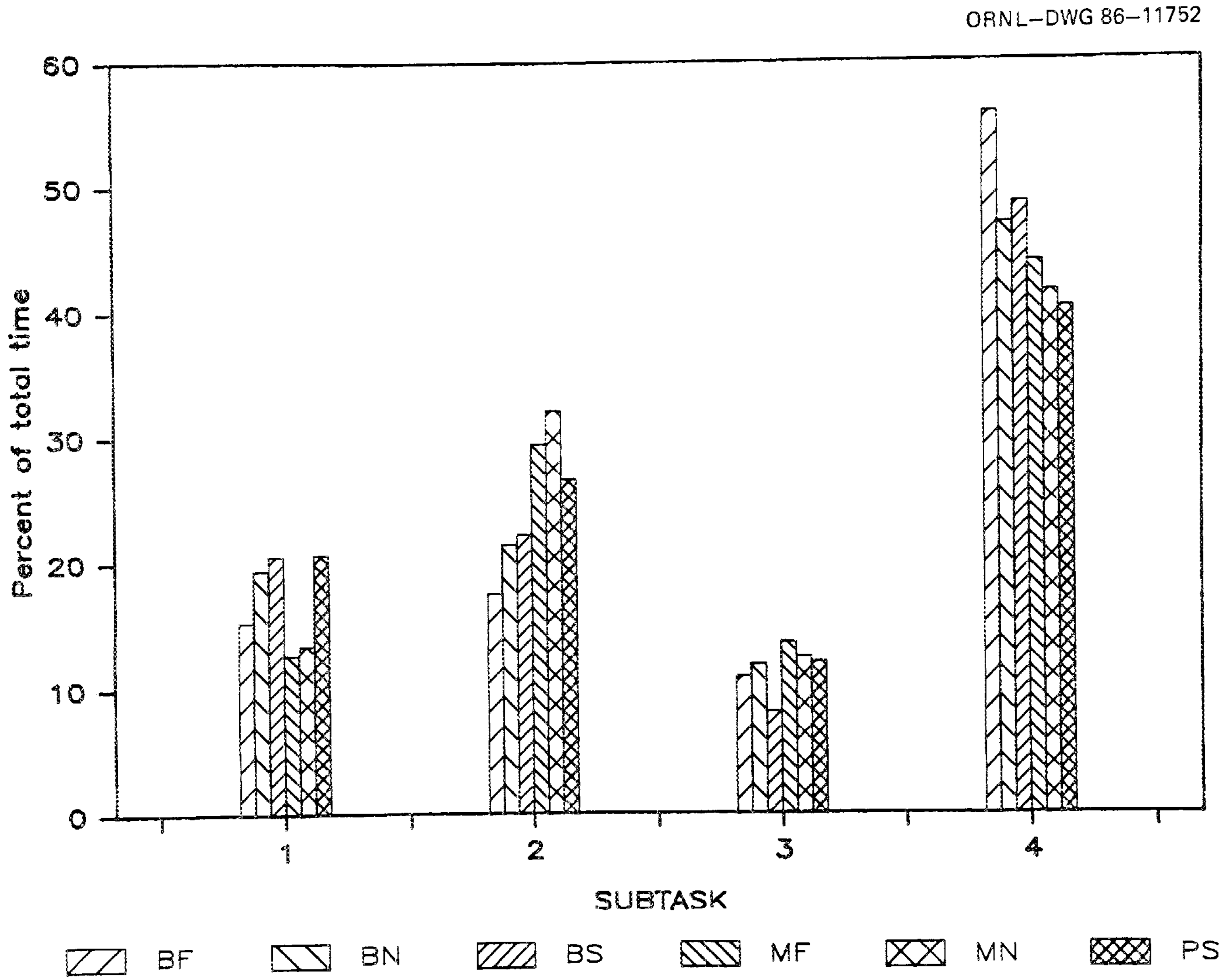

Fig. 28. Bolt-threading task.

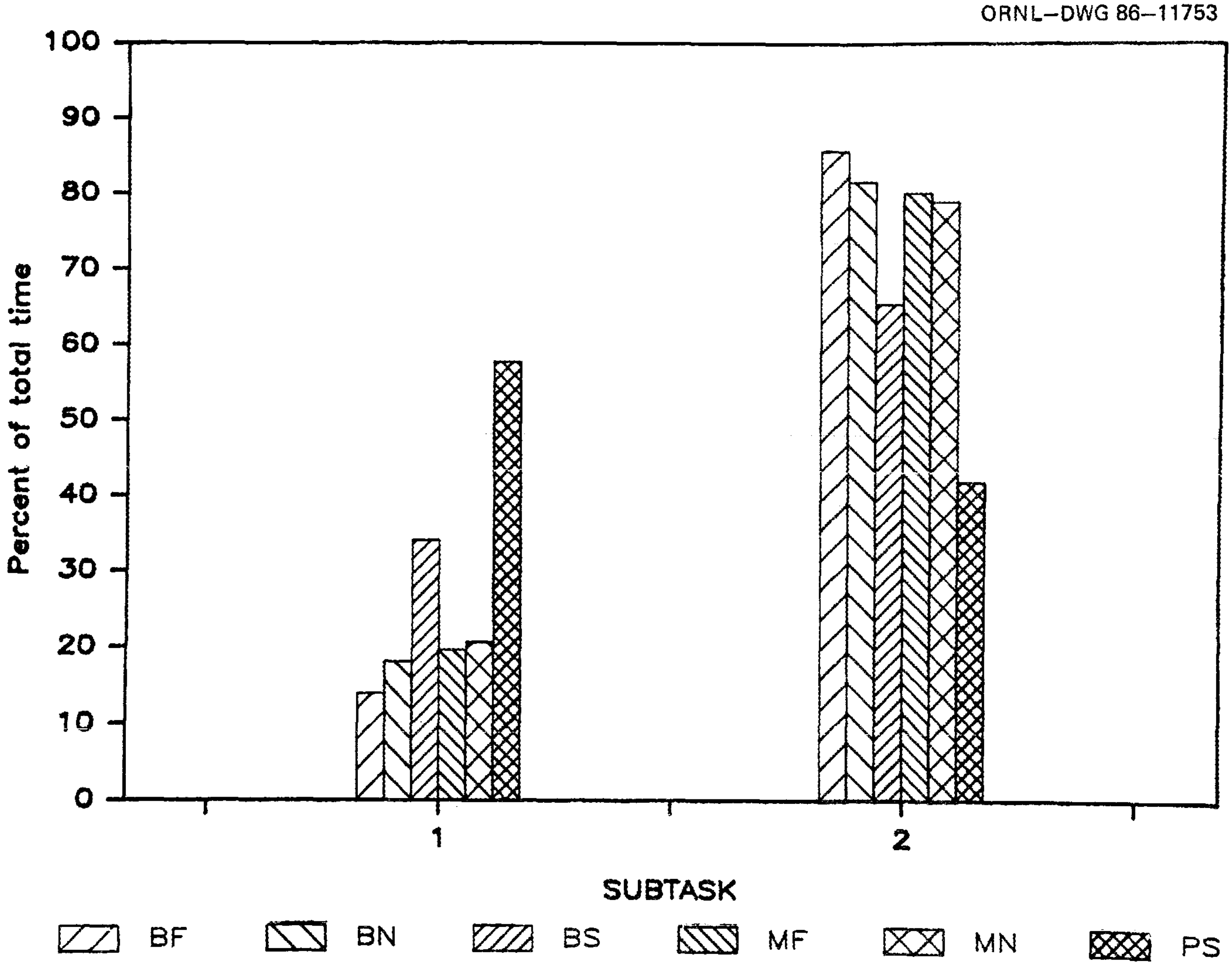

Fig. 29. Peg-in-hole task.

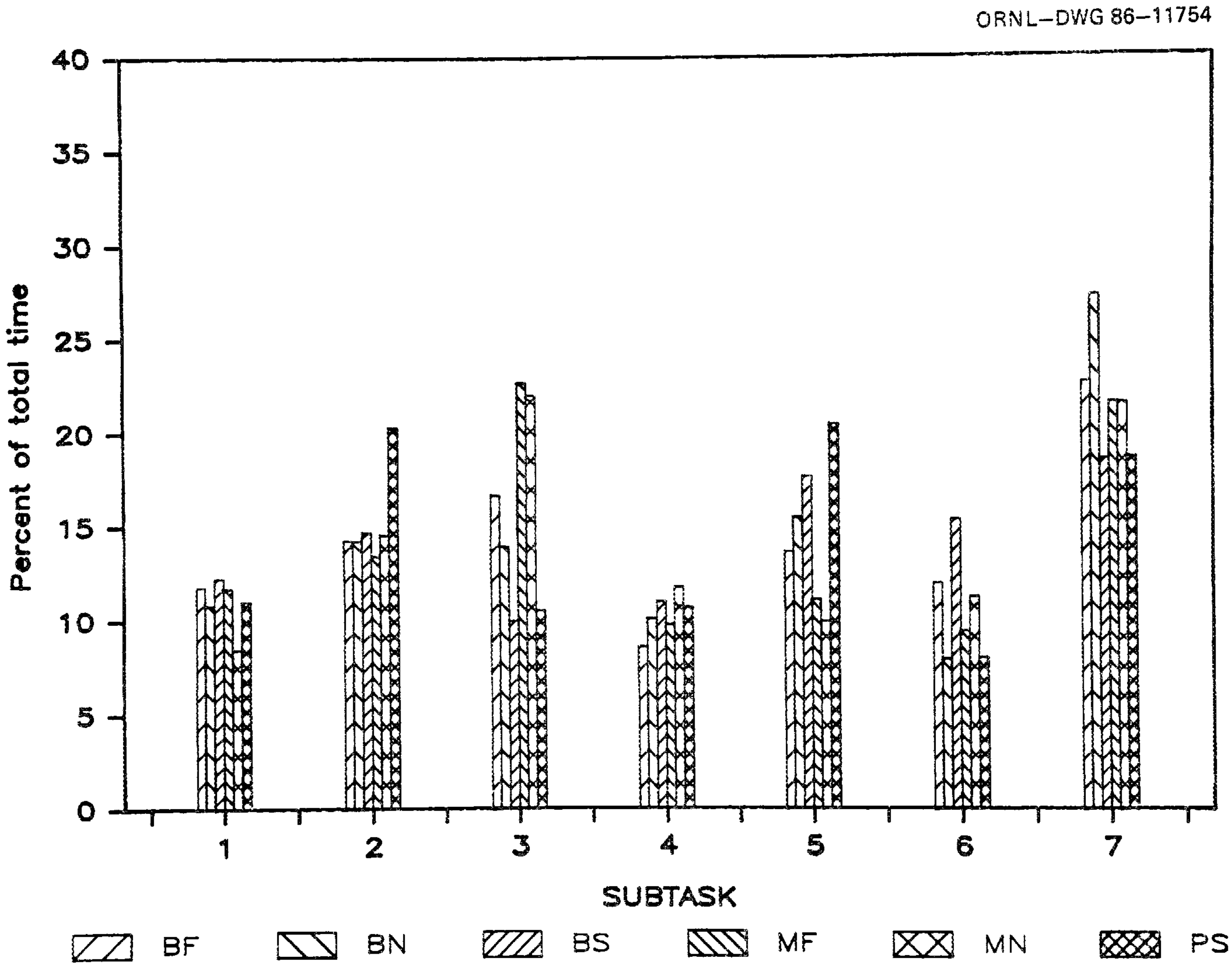

Fig. 30. Connector task.

5.4.6 Practice Effect

Figures 31 and 32 illustrate the effect of practice on task completion time and the total number of critical incidents per trial for the master/slave sessions in experiment 2. The plots are based on the average score per unit of four sessions, with each master/slave machine and mode used once in each unit. Both trends are significant (for task completion time, $t = 5.67$, $\alpha \leqslant 0.05$; for critical incidents per trial, $t = 5.5.1$, $\alpha \leqslant 0.05$), indicating that operators worked more quickly and more accurately as their experience with the tasks increased.

ORNL–DWG 86–11755

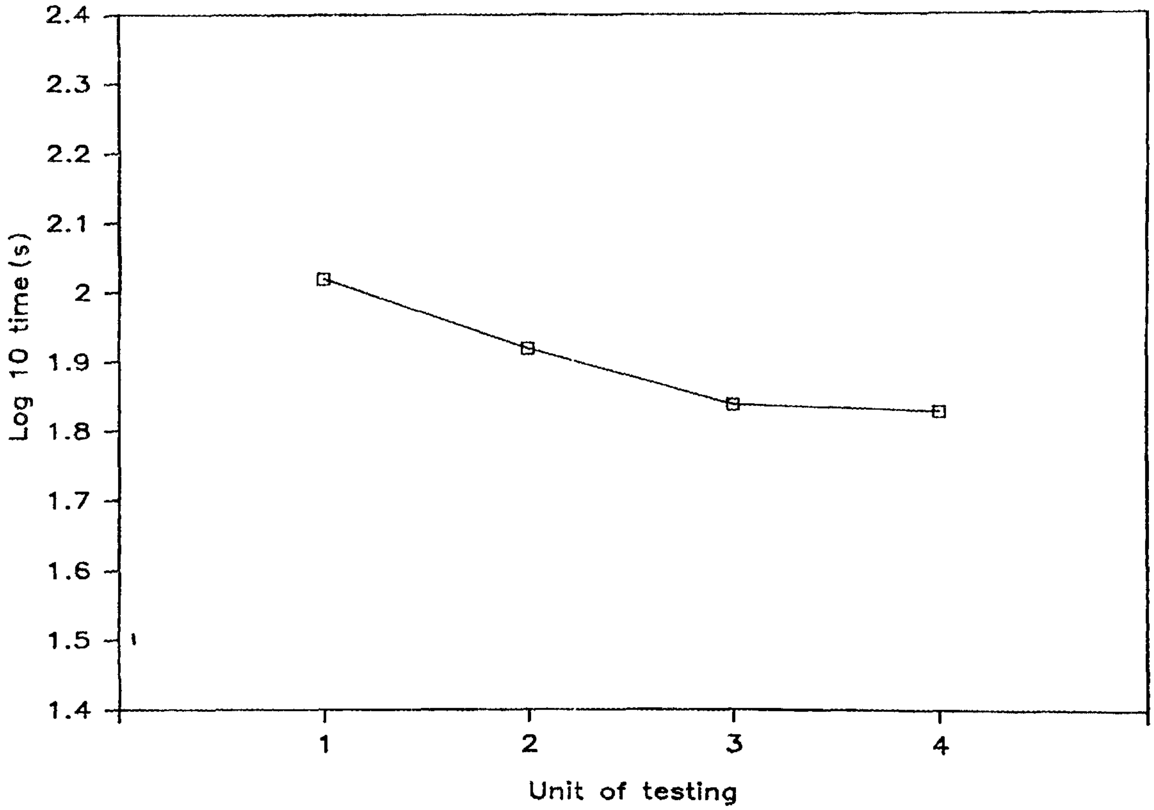

Fig. 31. Effect of practice on time.

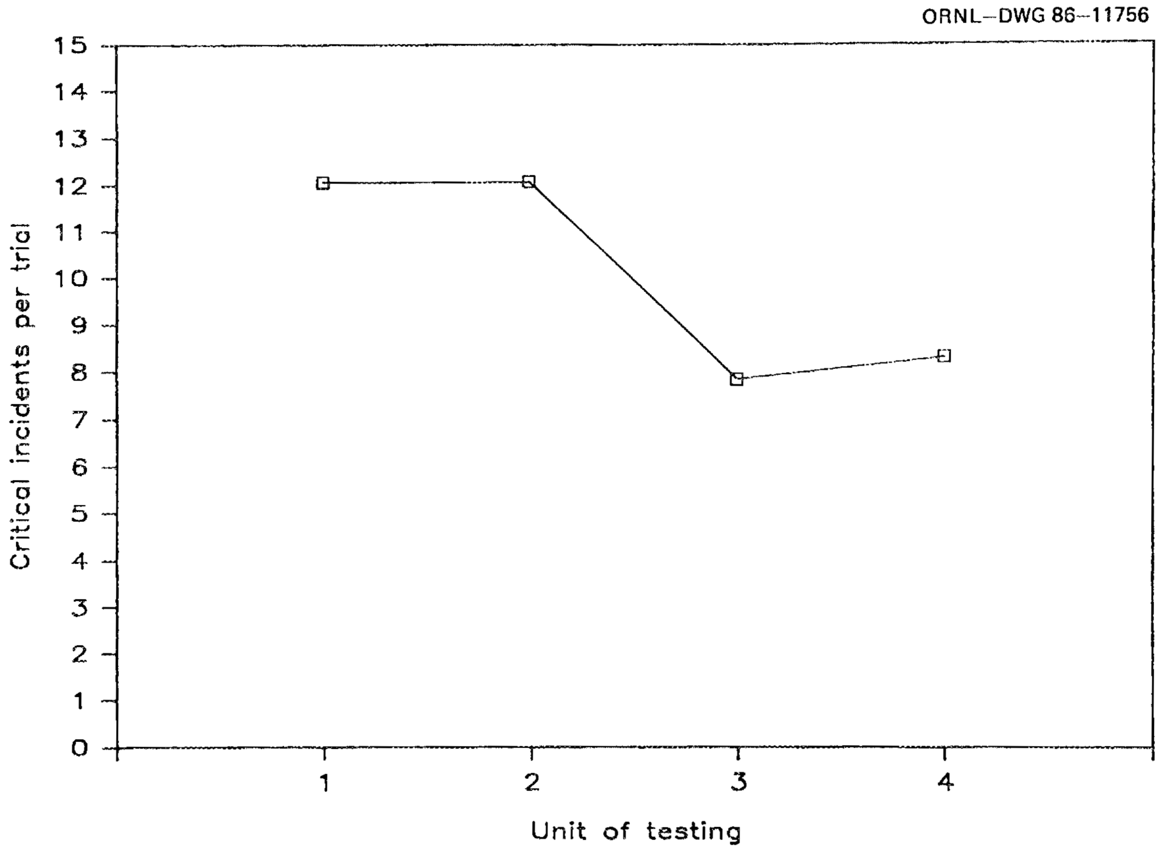

Fig. 32. Effect of practice.

6. EXPERIMENT 3

The results of experiment 3 were in agreement with the findings from experiments 1 and 2. Operators using the M-2 completed tasks more quickly and with fewer errors than they did when using the BILARM; force reflection degraded performance with the BILARM, but had no effect on or improved performance with the M-2; the BILARM performed better in master/slave mode than in switchbox mode. There were no differences between the BILARM in switchbox mode and the PaR arm in experiment 3.

6.1 TASK DESCRIPTION

The task used in experiment 3 was a mock-up of an instrument package that might be encountered in the maintenance of a fuel recycling facility (Fig. 33). The package consisted of a square steel framework attached to the horizontal surface of the test stand by two captive fasteners. A tubing jumper was connected to the top of the framework by a TRU connector, and another TRU connector attached to the jumper to the task board shelf. An electrical connector was attached to the top of the framework and to the vertical surface of the test stand. This task was selected to represent a typical full-component-replacement type operation; therefore, the results of experiments 1 and 2 (with task primitives) could be compared with the performance of a task more nearly that of a full remote maintenance operation.

The task started with the manipulator at a point 6 ft away from the mock-up. The operator moved the manipulator to the task area using the overhead transporter, picked up an impact wrench, disconnected the tubing jumper, set aside the wrench, disconnected the electrical connector, disconnected the framework from the horizontal mounting surface, picked up the framework, and returned to the start point. After successfully completing the disassembly of the mock-up, the operator returned the manipulator to the task site, aligned and seated the framework on the horizontal mounting surface, reconnected the captive fasteners, and reconnected the tubing jumper and the electrical connector. The operator completed the task by returning to the starting point.

The four operators who participated in experiment 3 were the same as those who took part in experiment 2.

Machine/mode combinations in experiment 3 were assigned orders in the same way that sequences were assigned in experiment 1. Table 19 lists the sequences for each operator, who completed one session with each machine mode/ combination.

ORNL–PHOTO 2486–85

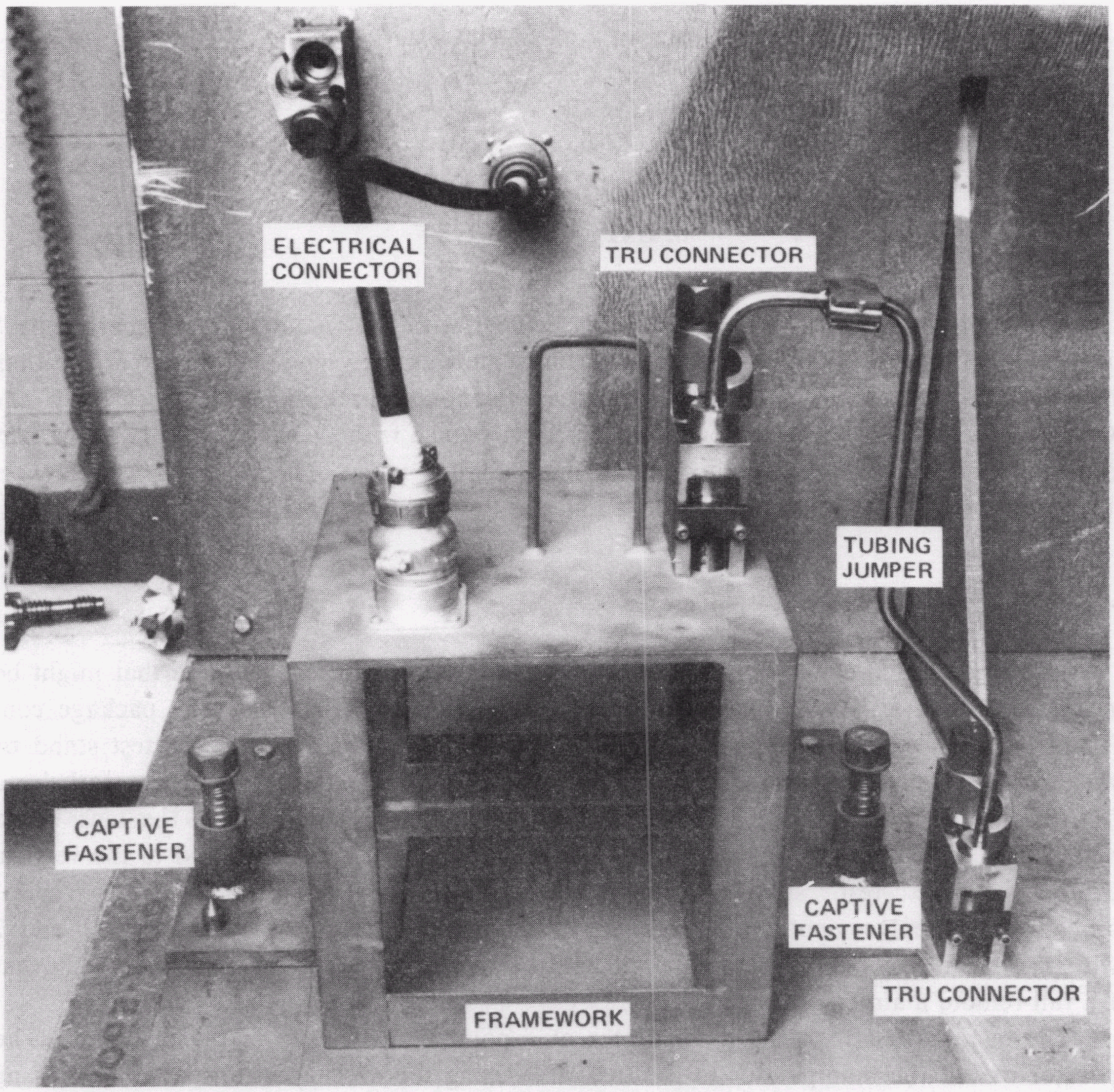

Fig. 33. Instrument package mock-up task, experiment 3.

Table 19. Testing sequence for each operator in experiment 3

Operator	Sequence					
1	BN[a]	MF[b]	MN[c]	BF[d]	BS[e]	PS[f]
2	BF	MF	MN	BN	BS	PS
3	MN	BF	BN	BF	BS	PS
4	MF	BF	BN	MN	BS	PS

[a]BILARM nonforce reflecting.
[b]M-2 force reflecting.
[c]M-2 nonforce reflecting.
[d]BILARM force reflecting.
[e]BILARM with switchbox.
[f]PaR manipulator.

6.2 PROCEDURES

Procedures were the same as for experiment 2, except that operators were allowed to complete as many repetitions of the task as they could (up to a limit of three repetitions) in a 1-h testing session. The dependent measures recorded for this experiment were the same as for experiment 2.

6.3 RESULTS

6.3.1 Machine Differences with Master/Slave Combinations

Operators using the M-2 manipulator completed the task in significantly less time than when using the BILARM ($t = 5.11$, $\alpha \leqslant 0.05$). They also committed significantly fewer critical incidents per trial ($t = 9.92$, $\alpha \leqslant 0.05$) with the M-2. Figure 34 illustrates the performance of all machine and mode combinations on the time-to-completion variable in this experiment, and Fig. 35 shows the performance on total critical incidents per trial.

ORNL–DWG 86–11757

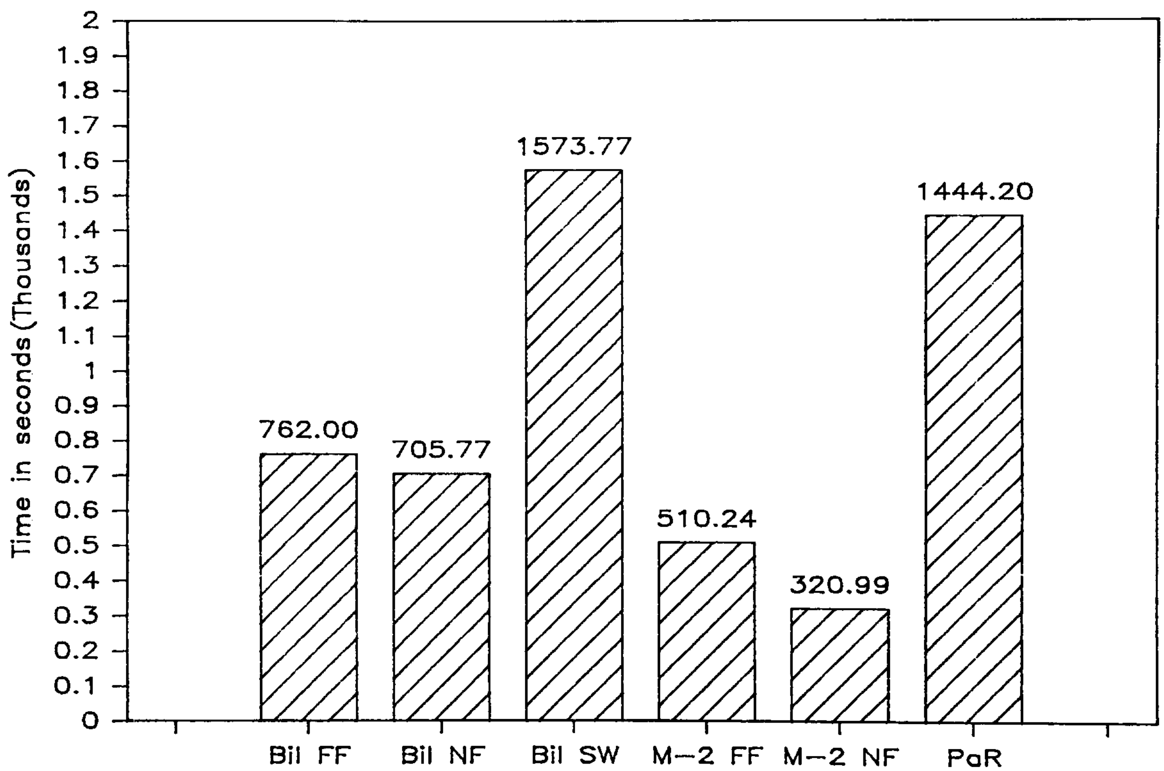

Fig. 34. Task time in experiment 3.

ORNL–DWG 86–11758

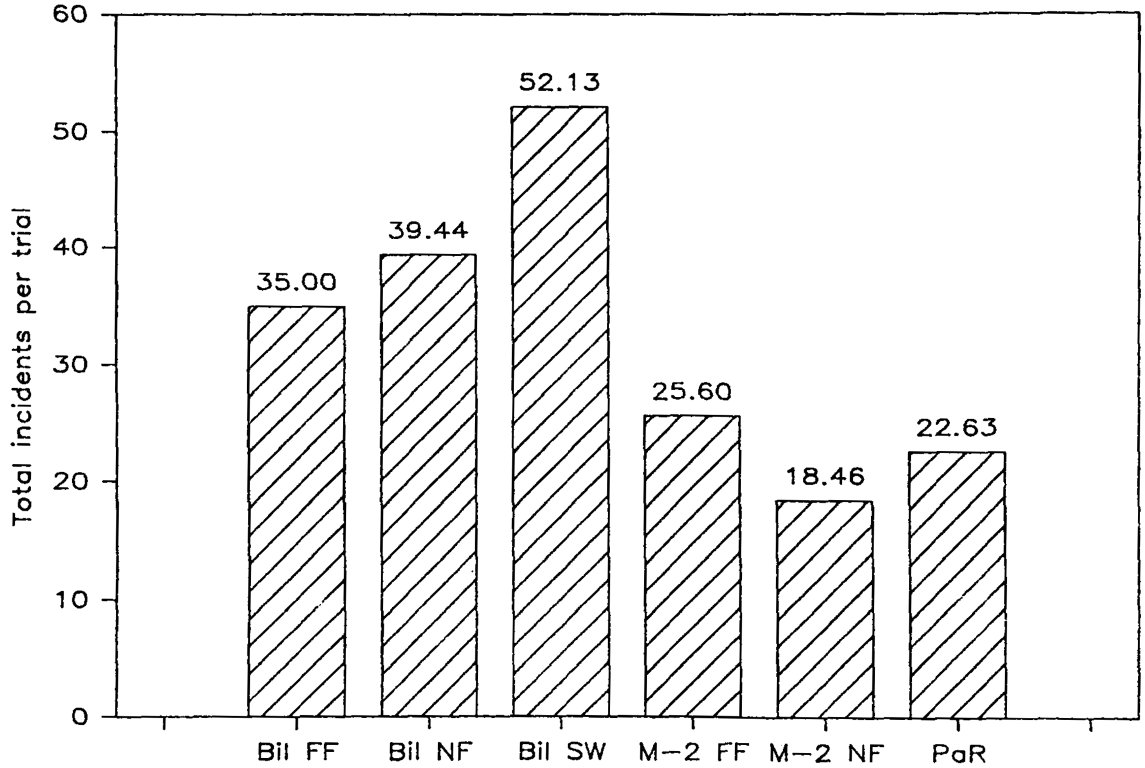

Fig. 35. Incidents per trial.

6.3.2 Force-Reflection Effects

There was no effect of force-reflection conditions on task completion time or total critical incidents in experiment 3. Although Figs. 34 and 35 seem to indicate that performance advantages existed for the nonforce-reflection condition for both the BILARM and the M-2, the differences were not statistically significant. The failure to find a statistically significant difference is because of differences in the way operators responded to force reflection on this task. Two of the operators completed the task in slightly less time with force reflection, whereas two other operators took considerably longer with force reflection. The difference in averages is reflected in the considerable decrease in performance in the force-reflection condition for the latter pair of operators.

6.3.3 Control Mode Differences Within the BILARM

Operators completed the task in significantly less time when using the master/slave mode than they did when using the switchbox controller ($t = 3.43$, $\alpha \leq 0.05$).

6.3.4 Machine Differences Within Switchbox Mode

There was no difference in the average time required to complete the task between the two switchbox-controlled manipulators. Operators using the BILARM committed significantly more critical incidents per trial than operators using the PaR manipulator ($t = 3.53$, $\alpha \leqslant 0.05$).

7. EXPERIMENT 4

The purpose of experiment 4 was to compare task performance using direct viewing with performance using television viewing. These data can also be used to construct manipulator direct-contact, task-completion-time ratios. This type of ratio has been used previously to compare manipulator system performance. However, previous ratios were based on manipulator/television to direct-contact/direct-viewing combinations. The television system degrades visual information; therefore, it should yield poorer performance regardless of the quality of the manipulator system. The results found here suggest that task completion times are roughly twice as long when using television viewing as they are when using direct viewing.

The tasks used in experiment 4 included all of the tasks included in experiments 1 through 3. Experiment 4 included all the operators who participated in experiments 2 through 3. During sessions in experiment 4, operators alternated between direct-viewing and television-viewing conditions to ensure that practice effects did not obscure differences between experimental conditions. The type of viewing system used first was determined randomly for each session.

7.1 PROCEDURES

7.1.1 Testing Sessions

Prior to performing tasks as scheduled in experiments 1 through 3, each operator completed the set of tasks manually. When the operator reported to the control room for testing, an observer met him and escorted him to the task site. The observer instructed the operator to perform the tasks in the order scheduled. The observer also told the operator the order of viewing conditions, that is, direct first and television second or television first and direct second. When the operator understood the order of task and viewing conditions, he positioned himself in front of the task area and waited for the observer to signal the start of the first trial. The observer initiated each trial by telling the operator to begin. The observer timed the task (using a hand-held stopwatch) and tallied the errors committed. At the end of each trial, the operator returned to the starting position and waited for the observer to tell him to begin the next task. During task performance, the observer monitored the operator to ensure that he used the correct viewing system.

7.1.2 Dependent Measure

Dependent measures were the logarithm to base 10 of task completion time and the tally of errors described in the procedures for experiment 1.

7.2 RESULTS

7.2.1 Differences Between Viewing Systems

The analysis of data from experiment 4 revealed a significant difference in time to completion and in errors between direct viewing and television viewing. Table 20 lists the averages of task time and errors for direct viewing and television viewing for each task.

Table 20. Average task completion time (in seconds) and errors for each viewing system and task in experiment 4

Task	Direct		Television	
	Time	Errors	Time	Errors
Impact wrench	3.00	0.224	5.56	1.828
Peg in hole	2.10	0.517	3.66	1.293
Motor mount	3.81	0.500	5.68	1.517
Bolt threading	2.99	0.180	3.81	0.597
Peg in hold (modified)	2.45	0.611	4.16	1.901
Connector	9.44	2.208	24.04	8.000
Instrument package mock-up	60.91	4.042	123.40	14.958

There were also significant differences between tasks in terms of the impact of the viewing system used, indicating that the performance of some tasks was more affected by the viewing system than other tasks. Figures 36 and 37 show this effect for task completion time and errors, respectively.

Table 21 lists average manipulator direct-contact, task-completion-time ratios for each manipulator and control mode combination for each task. The entries in the table are the ratios of the grand averages (averaging across operators and tasks) of task completion time while using manipulator systems (with television viewing) to the grand averages of task completion time when the tasks were completed manually (with direct viewing). Figure 38 illustrates these data.

ORNL--DWG 86-11759A

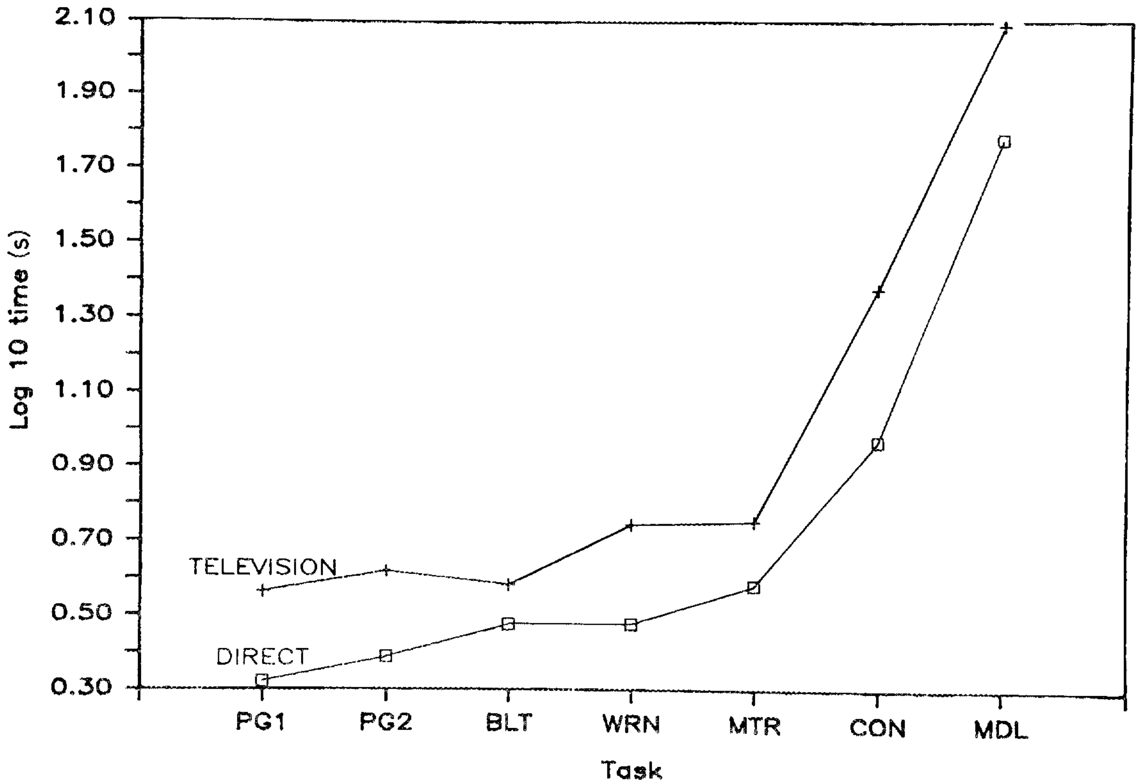

Fig. 36. Interaction, experiment 4.

ORNL–DWG 86-11760A

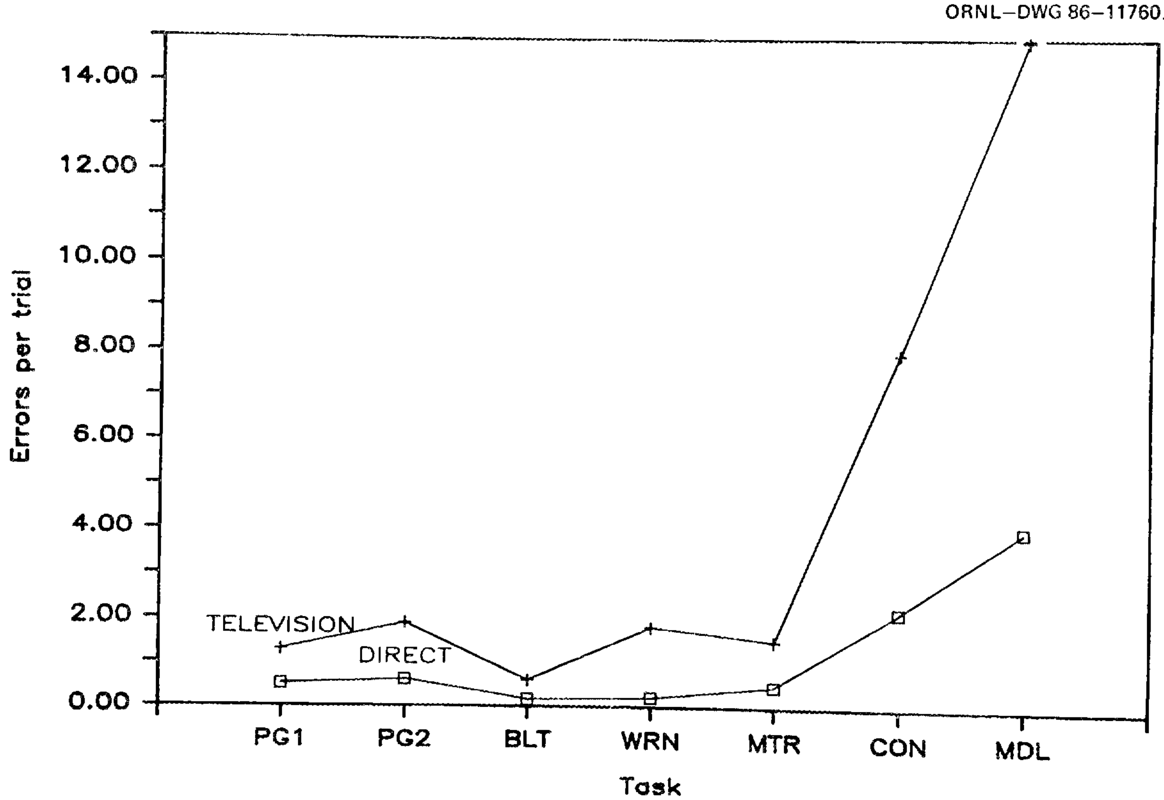

Fig. 37. Interaction, experiment 4.

Table 21. Task completion time ratios for each machine/mode combination and task

Task	Machine/Mode					
	BIL/FF	BIL/NF	BIL/SW	M2/FF	M2/NF	PaR
Impact wrench	12.61	12.47	38.98	7.28	6.64	42.35
Peg in hole	38.53	41.95	74.61	14.09	12.73	79.94
Motor mount	7.62	7.43	19.68	4.20	4.30	22.29
Bolt threading	58.66	35.25	91.91	13.98	12.98	39.29
Peg in hole (modified)	35.47	35.47	102.04	16.37	16.29	107.39
Connector	17.38	15.46	47.21	8.97	9.46	46.24
Instrument package mock-up	12.51	11.60	25.84	8.38	5.26	23.73
Average	26.11	22.80	57.18	10.47	9.60	52.09

ORNL–DWG 86–11761

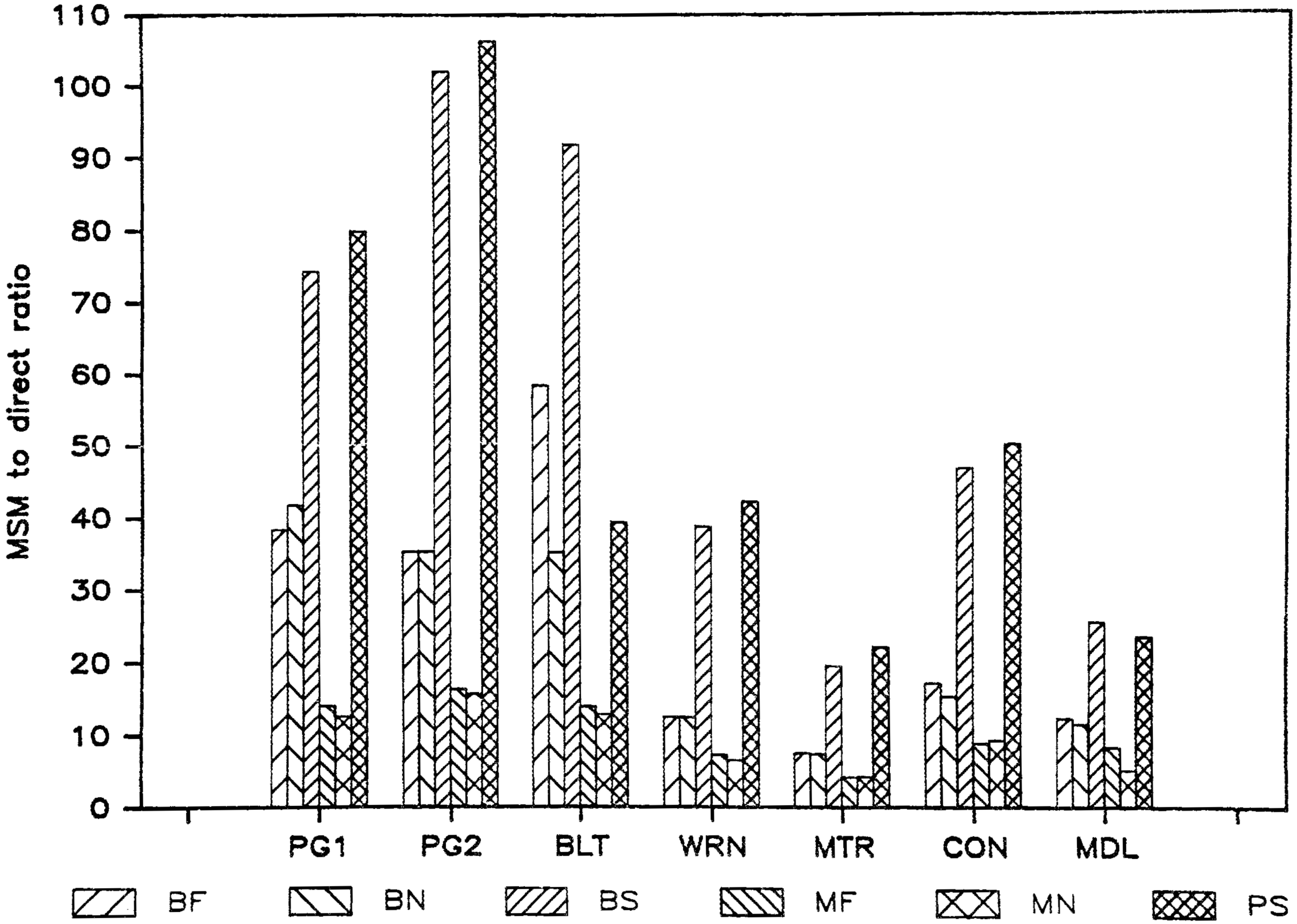

Fig. 38. Manipulator: direct contact.

8. GENERAL DISCUSSION

8.1 SUMMARY

In all stages of testing, the CRL Model M-2 manipulator had consistently lower times to completion than any other machine. The Meidensha BILARM 83A, operating in master/slave mode, was the second fastest machine (although occasionally it was no better than either of the switchbox-controlled systems). The BILARM in switchbox mode and the PaR 6000 manipulator were approximately equivalent in terms of performance variables during testing. The number of errors and critical incidents were similar, although differences between conditions were smaller than for task time. This suggests that this type of measure is a less sensitive criterion for comparisons among manipulator systems.

The effect of force reflection differed among manipulators. Force reflection adversely affected the performance of all tasks with the BILARM, but generally had the opposite effect for the M-2 (although the differences were not significant within the M-2). This reflects the lower fidelity and poorer kinematic arrangement for force reflection on the BILARM. Force reflection with the BILARM resulted in resistance to master control motions and failed to provide force information on a vertical axis (perpendicular to the yaw axis since no force reflection was provided by the elbow joint). These effects combined to make the BILARM a much less effective machine when its force reflection was engaged. On the other hand, there was no appreciable decrease in performance using force reflection with the M-2. In most cases force reflection had no effect on time to completion or time to subtask completion with the M-2 manipulator. Operators were able to complete tasks with fewer occurrences of some critical incidents with force reflection than without it. However, fpr the tasks used in this testing, the differences were not statistically significant.

These results should not be interpreted as evidence that force reflection is not beneficial in remote handling. These experiments failed to address the question of force reflection in a definitive fashion. The dependent measures collected were incomplete in terms of assessing the impact of force reflection, since the forces applied to task components were not measured. The testing program was designed to assess the strengths and weaknesses of the various manipulator systems used, which represent a fairly wide range of types of systems, from power manipulators to dexterous servomanipulators. To test this wide range of manipulators, it was necessary to restrict the range of tasks and conditions. For example, no tasks that required two arms to complete were included since that would have biased results in favor of the CRL M-2. The tasks were all designed to be robust enough to be performed by the PaR manipulator; more delicate tasks might have shown a difference in force reflection. In addition, testing conditions were not the type in which force reflection could be expected to have a dramatic effect. Trained operators were used in order to provide stable data for analysis of machine performance; force reflection is probably more beneficial to operators new to tasks or manipulators. Force reflection is helpful in cases where television does not provide adequate visual information, as is the case when a camera's lines of sight are obstructed by equipment at the task site; in these tests, cameras were positioned to provide good views of the remote site. Force reflection may also be helpful for trained operators performing a new task for the first time (which is the case for most maintenance jobs). The multirepetition approach used in this testing would not detect such an effect. The goals, conditions, and tasks of the testing program combined to make these data inadequate for making a final judgment on the efficacy of force reflection in remote handling. The data collected in manipulator comparative testing are useful for evaluating force reflection, but do not provide a complete picture of its effect.

The loss of information (compared to what an unencumbered human eye could detect) obtained through television transmission has been documented before. Experiment 4 demonstrates how pervasive this effect is; the tasks used were extremely simple, and yet the ratios of television viewing task time to direct viewing time are on the order of 2 to 1. With more complex tasks or with tasks that require more precise positioning of task components, the ratio could be much higher.

Another valuable point these data make has to do with comparisons of manipulators. The performance of these devices is often quantified by the ratio of remote performance to direct-contact performance. Remote handling devices have been compared using this type of figure of merit. Unfortunately, these ratios do not measure manipulator performance per se; instead, they measure the performance of the combination of manipulator system and viewing system. Unless the viewing systems of competing devices have been equated, any performance ratio confounds the effects of the handling systems with the effects of the viewing systems they use. In some cases it is appropriate to compare viewing system/manipulator system combinations. A comparison of direct-contact performance with master/slave manipulator performance should include the effects of the different viewing systems. The same can be said when comparing through-the-wall manipulators (which use shielded windows) and transporter-delivered manipulators. However, it is inappropriate to compare two manipulators that use different television systems, since performance differences may occur because of differences in viewing systems.

The Manipulator Comparative Testing Program contributed to the mission of the Fuel Recycle Division, ORNL, and to the mission of the PNC in Japan by demonstrating differences between remote handling systems, by making a preliminary investigation into the effects of force reflection on remote handling performance, and by providing a venue for the development and testing of methods for evaluating manipulator performance.

REFERENCES

1. R. G. O'Brien and M. K. Kaiser, "The MANOVA for Analyzing Repeated Measures Designs: An Extensive Primer," *Psychol. Bull.*, February 1985.

2. W. L. Hays, *Statistics for the Social Sciences*, 2nd ed. Holt, Rinehart, and Winston, New York, 1973, p. 424.

3. B. J. Winer, *Statistical Principles in Experimental Design*, McGraw-Hill, New York, 1971, p. 397.

Appendix

OVERVIEW OF STATISTICAL TESTS

This appendix explains some details of the statistical procedures used in analyzing data from the Manipulator Comparative Testing Program. It presents a brief overview of a complex topic. For more thorough coverage, consult the references listed at the end of the appendix or a professional statistician. The first paragraphs of the appendix define and explain some elementary statistical concepts; later paragraphs present details on the calculation of the t-test and of alpha.

The data from the Manipulator Comparative Testing Program were analyzed using a statistical test called Student's t-test, or simply the t-test. It is a common test for detecting differences between two averages. To understand the t-test, it is necessary to understand some basic statistical concepts, including the average, the standard deviation, and the normal distribution.

An average is a figure which summarizes a set of numbers. It is calculated by summing the set and dividing the sum by the number of members in the set. The average is the midpoint of the set. If the differences between the average and each member of the set are calculated, the sum of the differences between the average and numbers less than the average will be equal to the sum of differences between the average and numbers greater than the average. In any set of numbers, there is only one number that occupies this position in the middle of the set.

At this point a simple example will be helpful. Suppose we have the set of twelve numbers

(15 16 14 19 10 33 21 12 13 17 23 20)

The average of the set is equal to the sum of the members of the set divided by the number of members in the set, or

$$\frac{15 + 16 + 14 + 19 + 10 + 33 + 21 + 12 + 13 + 17 + 23 + 20)}{12}$$

The average of this set of numbers is 17.75. The subset of numbers less than 17.75 includes

(15 16 14 10 12 13 17),

and the subset greater than 17.75 includes

(19 35 21 23 20).

The sum of differences between the average and the numbers in the first subset (numbers less than the average) is equal to

$$(17.75-15) + (17.75-16) + (17.75-14) + (17.75-10) + (17.75-12) + (17.75-13) + (17.75-17) = -27.3$$

The sum of the second subset (numbers greater than the average) is

$$(17.75-19) + (17.75-35) + (17.75-21) + (17.75-23) + (17.75-20) = 27.3.$$

The sum of the differences in the first set and the differences in the second set is $-27.3 + 27.3 = 0$.

Another important statistic is the standard deviation. The standard deviation is a figure that expresses the amount of spread or dispersion of a set of numbers around their average. It is calculated by taking the square root of the average of the squared differences between each number in the set and the average of the set. The larger this number is, the greater the spread of the set around the average.

The calculations necessary for determining the standard deviation for the preceding example are given in Table A.1. In the table, x represents a member of the set and X represents the average. The first column shows the members of the set, the second column shows the difference between each member of the set and the average (17.75), and the third column presents the squared differences. The square root of the squared difference between a member of the set and the average is also called a deviation from the average. The square root of the average of the squared differences is the standard deviation.

The average and the standard deviation are useful because they allow quick interpretation of large sets of numbers. In the context of certain types of distributions of numbers, they become even more useful. The most important of these types of distributions is the normal distribution. Normally distributed sets of numbers have symmetrical, mono-modal frequency distributions (graphs of these distributions show one peak, and the regions on either side of the average are mirror images of each other), contain 68% of their members

Table A.1. Calculation of the standard deviation (SD)[a]

	x	x-X	$(x\text{-}X)^2$
	10	−7.75	60.06
	12	−5.75	33.06
	13	−4.75	22.56
	14	−3.75	14.06
	15	−2.75	7.56
	16	−1.75	3.06
	17	−0.75	0.56
	19	1.25	1.56
	20	2.25	5.06
	21	3.25	10.56
	23	5.25	27.56
	33	17.25	297.56
Sum	215	0.00	483.22

[a] $SD = \sqrt{483.22/12} = 6.35$

within one standard deviation of the average, contain 98% of their members within two standard deviations of the average, and contain 99% of their members within three standard deviations of the average.

Normal distributions are very useful, because the probability of any number being a member of a normally distributed set can be found using a z score. A z score is the difference between a number and the average of the normal distribution, expressed in standard deviation units. Specifically, $z = (x - X)/SD$, where x is the number in question, X is the average, and SD is the standard deviation of the distribution. The probability of any z score occurring in the distribution can be determined by looking it up in a table of z scores. For example, the number 33 in the example from Table A.1 has a z score of $z = (33-17.75)/6.35 = 2.40$. This is a high z score; if the set of numbers in the example were normally distributed, z scores this high would be very rare. In a normal distribution, z scores this high occur less than 1% of the time.

The z scores may be used to determine whether an observation belongs to the same set as other observations made under similar conditions. For example, the numbers in the example could be measurements of the time required to perform some task with a remote handling system. If an experimenter had reason to believe that the population of all possible observations followed a normal distribution, he could use the average and standard deviation calculated from his sample of measurements to estimate the values of the average and standard deviation of the population distribution. He could then use a z score to determine whether new observations could be from the same population. A measurement with a z score of 2.40 would have less than a 1% chance of coming from the same population of the other scores; the experimenter might conclude that this number if not part of the same set as the other numbers.

There is a known probability that the experimenter's conclusion is wrong. This probability is the same as the probability of such a z score occurring in the population, in this case $p(x = 33) = 0.0082$, according to a table of z scores found in Ref. 1. If the observation is in fact from the same population, the experimenter has made an error in judgement. In an experiment, the probability of this type of error occurring is called α (alpha). Usually, an experimenter decides on an acceptable probability of an error in judgement prior to data collection and makes up a decision rule based on this maximum acceptable probability of error. The decision rule takes the form "if the probability of an observation being part of the population is less than alpha, it will be considered significantly different from the population." While alpha may be arbitrarily set at any value, certain alpha levels have become standards among researchers in human factors and psychology. Alpha is very rarely set higher than $\alpha = 0.10$, and most frequently is set at $\alpha = 0.05$. High levels of alpha (0.06 to 0.10) are used in situations where measurement may be imprecise, or preliminary, exploratory experiments are being conducted. For research into large effects that can be measured very accurately, alpha may be set at $\alpha = 0.01$.

The t-test used in analysis of data from the Manipulator Comparative Testing Program is also based on characteristics on the normal curve, and it is quite similar to the z score. The t-test is an adaptation of the z score for small sets of numbers taken from normally distributed populations (although the test is fairly immune to departure of the population distribution from normality). In the Manipulator Comparative Testing Program, each operator contributed a score for each level of each factor under study. For example, each operator completed the peg-in-the-hole task with each force reflection level. A set of numbers was calculated by subtracting the scores for each operator for trials with force reflection from the score without it, and the average and standard deviation of the set of numbers were calculated. If there was no difference between the two conditions, the average difference should be close to zero. Another way to look at the problem is to say that if there is no real difference between the conditions, the set of numbers calculated by subtraction should be from a population of numbers which includes zero. To test this hypothesis, a t score is calculated for zero. The formula for the t score is

$$t = (x - X)/\sqrt{SD},$$

where x is zero, X is the average of the set, and SD is the standard deviation of the set. If performance in the two conditions is in fact different, the probability of obtaining the t score for zero should be less than the predetermined level of alpha, in this case 0.05. The probability of the t score occurring if performance under the two force reflection conditions is the same is found by consulting a table of values for t scores.

Several tables of t scores exist. The number of measurements in the set determines how closely the t distribution follows the normal distribution. The larger the size of the set, the smaller the deviation of the t distribution is from the normal distribution. The smaller the set, the greater the deviation is and the less sensitive the t-test becomes. Each t-test table is based on the degrees of freedom equal to the number of members of the set that may change without changing the average. In the example above, there are 12 members of the set; 11 of these could take on different values without changing the average, as long as the 12th member of the set was of a size which maintained the sum of the set. Therefore, the average in the example has 11 D.F. In the Manipulator Comparative Testing Program, each operator contributed one difference score (for example, the difference between trials with force reflection and trials without force reflection). Most of the number sets had 4 members, and therefore had 3 D.F.

This overview of the statistical procedures used in the Manipulator Comparative Testing Program has been necessarily brief. For a more detailed description of the philosophy and mechanics of the t-test and other statistical tests, consult these references:

W. L. Hayes, *Statistics for the Social Sciences*, 2nd ed., Holt, Rinehart, and Winston, New York, 1973.

L. Blank, *Statistical Procedures for Engineering, Management, and Science*, McGraw-Hill, New York, 1980.

W. J. Diamond, *Practical Experimental Designs*, Lifetime Learning Publications, Belmont, Calif., 1981.

ORNL/TM-10109
Dist. Category UC-86

INTERNAL DISTRIBUTION

1. D. E. Benker
2. J. O. Blomeke
3. E. C. Bradley
4–6. W. D. Burch
7–9. T. W. Burgess
10. R. H. Chapman
11. E. D. Collins
12. S. F. DeMuth
13–17. J. V. Draper
18. R. D. Ehrlich
19. J. H. Evans
20. M. J. Feldman
21. N. R. Grant
22. W. S. Groenier
23. W. R. Hamel
24. D. C. Hampson
25–26. S. J. Handel
27. T. L. Hebble
28–32. J. N. Herndon
33. L. J. King
34. E. H. Krieg, Jr.
35. C. T. Kring
36. D. P. Kuban
37. R. E. Leuze
38–43. S. A. Meacham
44. C. F. Metz
45. D. R. Moser
46. F. L. Peishel
47. W. W. Pitt
48. D. P. Reid
49. J. C. Rowe
50. S. L. Schrock
51. J. H. Shaffer
52–54. R. W. Sharpe
55. G. E. Smith
56. B. B. Spencer
57. J. G. Stradley
58–59. E. Sundstrom
60. C. R. Weisben
61–62. T. J. Williams
63. O. O. Yarbro
64. H. R. Yook
65–66. Laboratory Records
67. Laboratory Records, ORNL-RC
68. ORNL Patent Section

EXTERNAL DISTRIBUTION

69. D. E. Bailey, Director, Division of Fuels and Reprocessing, Office of Facilities, Fuel Cycle, and Test Programs, U.S. Department of Energy, Washington, DC 20545
70. F. P. Baranowski, 1110 Dapple Grey Court, Great Falls, VA 22066
71. S. J. Beard, Vice President, Marketing and Uranium Operations, Exxon Nuclear Company, Inc., 600 108th Avenue, N.E., C-00777, Bellevue, WA 98009
72. A. Bejcay, Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109
73. J. Byrd, E. I. du Pont de Nemours and Company, Savannah River Laboratory, Aiken, SC 29801
74. M. M. Clarke, Rockwell International, 12214 Lakewood Blvd., Downey, CA 90241
75. Y. Fujita, PNC Tokai Works, 4-33 Muramatsu Tokai-Mura, Naka-gun, Ibaraki-Ken Japan
76. L. Jenkins, NASA, Johnson Space Center, Houston, TX 77058
77. M. Maeda, PNC Tokai Works, 4-33 Muramatsu Tokai-Mura, Naka-gun, Ibaraki-Ken Japan

78. H. L. Martin, Telerobotics International Inc., 8410 Oak Ridge Highway, Knoxville, TN 37931
79. A. J. Meintell, NASA, Langley Research Center, Hampton, VI 23665
80. J. Oberright, NASA, Goddard Space Flight Center, Greenbelt, MD 20771
81. M. J. Ohanian, Associate Dean for Research, College of Engineering, 300 Weil Hall, University of Florida, Gainesville, FL 32611
82. J. E. Pennington, NASA, Langley Research Center, Hampton, VI 23665
83. R. Pepper, Naval Ocean Systems Center, Ocean Engineering Division, P.O. Box 997, Kailua, HI 96734
84. J. F. Proctor, Senior Technical Specialist, E. I. du Pont de Nemours and Company, Savannah River Laboratory, Aiken, SC 29801
85. E. C. Smith, NASA, Marshall Space Flight Center, Huntsville, AL 35812

86–96. H. Takeda, General Manager, Planning Section, Reprocessing Development Division, Power Reactor and Nuclear Fuel Development Corporation, 9-13, 1-Chome, Akasaka, Minato-ku, Tokyo Japan

97. Office of Assistant Manager for Energy Research and Development, DOE-ORO, Oak Ridge, TN 37831

98–204. Given distribution as shown in TIC-4500 under UC-86, Consolidated Fuel Reprocessing Category