MARTIN MARIETTA ENERGY SYSTEMS LIBRARIES
3 4456 0155142 3

ORNL/TM-10303

ornl

OAK RIDGE
NATIONAL
LABORATORY

MARTIN MARIETTA

Evaluation of High-Definition Television for Remote Task Performance

J. V. Draper
Y. Fujita
J. N. Herndon

OAK RIDGE NATIONAL LABORATORY
CENTRAL RESEARCH LIBRARY
CIRCULATION SECTION
4500N ROOM 175
LIBRARY LOAN COPY
DO NOT TRANSFER TO ANOTHER PERSON
If you wish someone else to see this report, send in name with report and the library will arrange a loan.

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: A03 Microfiche A01

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

ORNL/TM-10303
Dist. Category UC-86

Consolidated Fuel Reprocessing Program

EVALUATION OF HIGH-DEFINITION TELEVISION FOR REMOTE TASK PERFORMANCE

J. V. Draper
Human Machine Interfaces, Incorporated

Y. Fujita
Power Reactor and Nuclear Fuel
Development Corporation (Japan)

J. N. Herndon
Fuel Recycle Division
Oak Ridge National Laboratory

Date Published: April 1987

Prepared for the
Office of Facilities, Fuel Cycle,
and Test Programs

Prepared by the
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

MARTIN MARIETTA ENERGY SYSTEMS LIBRARIES
3 4456 0155142 3

CONTENTS

ABSTRACT v

1. INTRODUCTION 1

2. EXPERIMENT 1: DEPTH PERCEPTION 3
   - 2.1 Experimental Design 3
     - 2.1.1 Experimental Task 3
     - 2.1.2 Dependent Variables 5
     - 2.1.3 Operators 5
     - 2.1.4 Procedures 5
   - 2.2 Results 5
     - 2.2.1 Data from 50-mm Separation Trials 7
     - 2.2.2 Data from 100-mm Separation Trials 8
   - 2.3 Conclusions 8

3. EXPERIMENT 2: REMOTE HANDLING 11
   - 3.1 Experimental Design 11
     - 3.1.1 Experimental Task 11
     - 3.1.2 Dependent Variables 12
     - 3.1.3 Operators 12
     - 3.1.4 Procedures 12
   - 3.2 Results 13
   - 3.3 Conclusions 13

4. EXPERIMENT 3: VISUAL INSPECTION 17
   - 4.1 Experimental Design 17
     - 4.1.1 Television System 17
     - 4.1.2 Experimental Task 17
     - 4.1.3 Participants 19
     - 4.1.4 Procedure 19
   - 4.2 Analysis of Data 19

5. DISCUSSION 23

REFERENCES 25

ABSTRACT

High-definition television (HDTV) transmits a video image with more than twice the number (1125 for HDTV to 525 for standard-resolution TV) of horizontal scan lines that standard-resolution TV provides. The improvement in picture quality (compared to standard-resolution TV) that the extra scan lines provide is impressive. Objects in the HDTV picture have more sharply defined edges, better contrast, and more accurate reproduction of shading and color patterns than do those in the standard-resolution TV picture. Because the TV viewing system is a key component for teleoperator performance, an improvement in TV picture quality could mean an improvement in the speed and accuracy with which teleoperators perform tasks.

This report describes three experiments designed to evaluate the impact of HDTV on the performance of typical remote tasks. The performance of HDTV was compared to that of standard-resolution, monochromatic TV and standard-resolution, stereoscopic, monochromatic TV in the context of judgment of depth in a televised scene, visual inspection of an object, and performance of a typical remote handling task.

The results of the three experiments show that in some areas HDTV can lead to improvement in teleoperator performance. The HDTV is superior to monoscopic, monochromatic, standard-resolution TV and to stereoscopic TV for remote inspection tasks; it is less proficient than stereo TV for interpretation of the distance between objects in the remote area. The HDTV leads to a lower rate of errors committed during remote handling tasks, but it does not reduce the amount of time necessary to perform the tasks.

While the depth perception task showed a performance advantage for stereo TV, it may be that the task emphasized the type of depth information provided by stereo TV. Stereo TV provides cues from retinal disparity. Other cues are derived by comparing perspective, size, texture, and patterns of light and shadow. These cues may be enhanced by the higher resolution provided by HDTV. Information from this type of cue was present to a very limited extent in the depth perception task used in this experiment. The relative performance of the TV systems in the depth perception experiment may not be directly translatable to all situations which require depth judgments.

On the remote handling task, HDTV allowed the operators to work more accurately than with the other systems, with no sacrifice of task-completion time. The sharper edges provided by the ultra-high-resolution HDTV systems allowed the operators to make more accurate judgments of the relative positions of the task components. It helped them avoid accidental contacts in the remote area and to make more positive contact with items to be grasped.

Observers inspecting a small object for a flaw were more accurate with HDTV than with either of the standard-resolution systems. High resolution is critical for detection of small-scale flaws of the type in the experiment (a scratch on a glass bottle).

These experiments provided an evaluation of HDTV television which is fairly wide in scope. The tasks selected were representative of a variety of the tasks that must be routinely performed to remotely maintain a nuclear fuel reprocessing facility.

1. INTRODUCTION

High-definition television (HDTV) transmits a video image with more than twice the number of horizontal scan lines as standard-resolution television (1125 from HDTV to 525 for standard-resolution TV). The added scan lines dramatically improve the resolution of images on the TV monitor. The resulting improvement in picture quality (compared to standard-resolution television) is impressive. Objects in the HDTV picture have more sharply defined edges, better contrast, and more accurate reproduction of shading and color patterns than do those in the standard-resolution TV picture. To the casual observer, the HDTV picture provides a greater sensation of depth.

Humans are primarily visually oriented.[1] Operators of remote handling equipment depend on TV systems for visual information they use to perform remote tasks. Because the TV viewing system is a key component for teleoperator performance, an improvement in the quality of TV pictures could mean an improvement in the speed and accuracy with which teleoperators perform their tasks. This report describes experiments designed to evaluate the impact of HDTV on the performance of typical remote tasks.

The experiments described in this report compared the performance of operators using HDTV with their performance while using other TV systems. The experiments included four TV systems: (1) high-definition, color TV (abbreviated HDCL); (2) high-definition, monochromatic TV (HDMO); (3) standard-resolution, monochromatic TV (SRMT); and (4) standard-resolution, stereoscopic TV (SRST). The stereoscopic TV was also monochromatic. The stereo system accomplished stereoscopy by displaying two cross-polarized images, one reflected by a half-silvered mirror and one seen through the mirror. Observers wore a pair of glasses with cross-polarized lenses so that the left eye received only the view from the left-hand camera and the right eye received only the view from the right-hand camera. A pair of cameras mounted approximately 5 in. apart with aiming lines converged on a point in the remote area provided the TV images.

Stereo TV restores some of the binocular depth cues in televised scenes that monoscopic television removes by displaying the image on a flat surface. The most important of these binocular cues is retinal disparity. In unaided vision, the images impinging upon the retinae are from slightly different perspectives. The difference in the images provides information about the distance to the object creating the image. In standard-monoscopic television, the images striking the retinas are identical because both eyes are viewing the same image on the TV monitor. Therefore, all cues to distance that might be provided by retinal disparity are lost. Stereo TV restores retinal disparity by providing a different view of the televised scene to each eye.

Three tasks were completed in the experiments: a depth perception test, a visual inspection task, and a small-scale remote handling task. This combination is representative of the types of visual tasks teleoperators perform while doing remote maintenance. Many

different tasks require alignment of objects in three dimensions. With the loss of depth cues compared to those available with the naked eye, judgment of distances from the camera's point of view is extremely difficult. A test of depth perception evaluated the ability of the sharper HDTV image to provide depth information. The depth perception test required operators to adjust the apparent distance to a section of polyvinyl chloride (PVC) pipe so that the pipe section was the same distance away as a second, stationary section of pipe. Measurement of the alignment error (the difference in position of the two pipe sections along the depth axis) of the pipe sections allowed comparison of performance in a situation requiring accurate perception of distances from televised scenes.

Remote maintenance activities in future nuclear fuel reprocessing plants will include inspection of items such as welds, pipes, and containers. While some of these tasks may be performed by robots, human inspectors will continue to have a role in inspections. Therefore, a task assessing remote inspection performance was included in the experiments. R. A. Erickson reports that the visibility of an object displayed on TV is affected by the number of scan lines which composed the image.[2] Because HDTV images are composed of more scan lines than images displayed by conventional TV, inspectors using HDTV should be able to spot objects that occupy a smaller area within the televised image than inspectors using conventional TV are able to detect. Inspectors using HDTV should be more accurate at detecting objects that are the same size on HDTV and standard-resolution TV. The inspection task included in these experiments involved checking a 10-mL sample bottle from an automated sampling system for a flaw on its exterior surface.

Remote handling tasks require dexterous manipulation using television as the primary sensory mode. Visual information (including information pertaining to depth and alignment) is crucial for good remote handling performance. Precise alignment in three-dimensional space is a task that is particularly dependent upon accurate visual information. Therefore, the remote handling task required assembly of four military-type electrical connectors. Military connectors are round, multiprong, plug-in connectors. Without precise alignment of the male and female portions of the connectors, coupling is impossible, and damage to the connectors is likely. This task required operators to make a critical visual alignment and to judge depth to complete it. Operators using HDTV should be able to connect a series of military connectors in less time and with fewer errors of alignment then operators using standard television.

Three separate experiments were conducted. In the first, operators completed the depth perception test, and alignment error (the distance from the front of the stationary pipe to the front of the movable pipe) was measured. In the second, operators performed the small-scale remote handling task, and task completion time and errors were measured. In the third, operators inspected the remote sample bottle using TV, and the accuracy of fault detection was measured.

2. EXPERIMENT 1: DEPTH PERCEPTION

2.1 EXPERIMENTAL DESIGN

2.1.1 Experimental Task

Figure 1 is a photograph of the apparatus used in this experiment. The operators sat directly in front of a TV monitor that displayed an image of the apparatus. The image on the TV monitor showed the upper half of the two pipes. No other portion of the apparatus was visible during testing.

The task involved moving one of the sections of gray PVC pipe so that the front surface of the pipe seemed to be the same distance away as a second, stationary section of

ORNL–PHOTO 8483–85

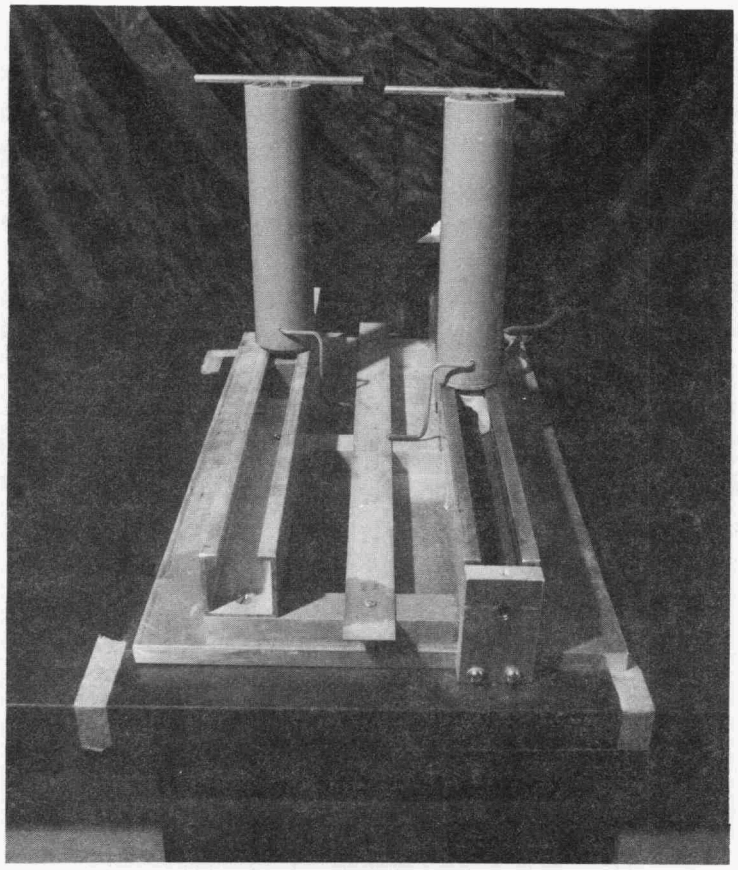

Fig. 1. Depth perception task.

PVC pipe visible in the televised scene. The pipe sections were mounted at the same height and about 3 in. apart. The operators moved the section of pipe by means of a toggle switch that controlled an electric motor. The motor rotated a worm gear attached to the bottom of the pipe section. Pushing the toggle toward the screen increased the distance between the pipe section and the camera, and pulling the toggle toward the operator decreased the distance between the pipe section and the camera. The movement of the pipe was quite slow (less than 1 cm/s), and the moving section stopped immediately on release of the toggle switch.

2.1.2 Dependent Variables

An experimenter who supervised testing sessions recorded the actual separation along the depth axis (the axis described by the line of sight of the camera) between the movable pipe section and the stationary pipe section at the end of each trial. The measurement was made using a metric scale attached to the apparatus. The measurement was made in millimeters.

2.1.3 Operators

Four operators participated in the experiment. All the operators were between the ages of 21 and 28, and each reported having 20/20 vision or vision corrected to 20/20. Each completed four 90-min testing sessions, one with each of the four TV systems. In each session, the operators completed 80 trials, 10 of which were practice sessions and are not included in the data. The order of TV system presentation was counterbalanced across operators so that no system was presented more frequently than any other at the start of testing or at the end of testing. This procedure prevented the effect of operator practice from being confounded with the effects of the television systems under study. For example, if any system were consistently presented at the start of testing, its performance could be expected to be worse than that of other systems. Operators will be better at the tasks at the end of testing than at the beginning of testing because of practice. The results of an experiment conducted without counterbalancing the effect of practice could be misleading. Table 1 lists the order of presentation of the TV systems for each operator.

Table 1. Order of TV system presentation, depth perception experiment

	Testing session			
Observer	1			
1	HDCL	SRST	HDMO	SRMT
2	HDMO	SRMT	HDCL	SRST
3	HDMO	SRMT	SRST	HDCL
4	SRST	HDCL	SRMT	HDMO

2.1.4 Procedures

Observers reported to the testing area and were seated at a long table on which the monitors for the four TV systems rested. The operators were given a control box with the three-position toggle switch (which controlled the direction of motion of the movable pipe section) mounted on its upper surface. The experimenter was seated behind the operator, next to the pipe section apparatus and the TV cameras. During testing, the operator was unable to see the cameras or the pipe-section apparatus. During a testing session, only the TV system used in that session was powered. The operator sat directly in front of the monitor for that system.

Before the start of each trial, the experimenter directed the operator to look away from the TV monitor and then placed the movable pipe section into one of four possible positions relative to the stationary pipe section, according to a random sequence of positions developed prior to the testing session. The positions the movable pipe section might occupy at the start of a trial were 100 mm in front of the stationary section, 50 mm in front of it, 50 mm behind it, and 100 mm behind it. To start each trial, the experimenter indicated to the operator that the operator could begin, the operator looked at the TV monitor, and the trial started. The trial ended when the operator indicated to the experimenter that he had completed positioning the movable pipe section. At that point, the experimenter measured the separation between the pipe sections.

2.2 RESULTS

The data were treated by the method of average error.[3] Using this method, the accuracy of alignment of the two pipe sections is measured by the standard deviation of the distribution of alignment errors. The smaller the standard deviation of the distribution of alignment errors, the more sensitive the observer is to differences in position, and the more precise judgments of position are possible. By making a series of judgments with each TV system, it was possible to calculate a standard deviation for each system. Comparison of the standard deviations indicates which TV system allows the most precise positioning of the movable pipe section, and, therefore, which provides the most accurate depth information.

A preliminary repeated-measures analysis of variance[4] (ANOVA) was conducted to determine whether the data taken from the four different starting positions (movable pipe section 100 or 50 mm from the stationary pipe section, in front of or behind the stationary pipe section) of the movable pipe section could be averaged together. Analysis of variance determines whether a set of averages contains two or more averages different enough to be considered statistically different. Statistically significant differences are large enough to be very unlikely to occur by chance; one may conclude that the observations are from different populations of observations. If the observations are, in fact, from the same population, this is an inappropriate conclusion. 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 judgment prior to data collection and makes up a decision-rule

based on this maximum acceptable probability of an error. The decision-rule takes the following 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 figure, 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; where preliminary, exploratory experiments are being conducted; or where it is more important to detect any differences than it is to avoid errors. For research into large effects that can be measured very accurately, alpha may be set at $\alpha = 0.01$.

The ANOVA makes a test called the F test (or simply F), which estimates the probability that the differences observed could have occurred by chance. The F test is a ratio of the variance of the averages around the overall average to an estimate of error (random) variance observed in the data. The first step in finding F is calculating sums of squares (SS), the sum of the squared differences between the averages and the overall mean. The sums of squares are divided by their degrees of freedom to calculate estimates of variance called mean squares (MS). The sensitivity of the F test to differences (that is, the likelihood that F will discover significant differences if they do indeed exist) is affected by the degrees of freedom of the F test. The degrees of freedom of F are the number of experimental units that contribute to the averages being tested and which are free to vary without changing the overall average. For example, to test if four operators have significant differences between their scores, ANOVA would test four averages. Three of these could take on different values without changing the overall average; therefore the test has 3 degrees of freedom (D.F.). The more degrees of freedom an F test has, the greater its sensitivity. The F test is a ratio of the mean square for the averages of interest and the mean square for error.

The probability estimate for the F test is called α (alpha). Alpha for a given F test is found by comparing the figure calculated for F to figures found in standard tables of F with the degrees of freedom for the numerator mean square and denominator mean square degrees of freedom. Differences are normally considered significantly different if α is equal to or smaller than 0.05.

In this analysis, a significant difference between initial pipe section separation conditions would mean that the data should be treated separately. In fact, as Table 2 shows, there was a significant difference between the averages of alignment error for the four initial positions ($F[3, 9] = 58.19, < 0.01$). Note that some rows in Table 2 do not show F tests; these are factors for which no test exists in the model used. They represent errors of measurement, and some of them are used as denominators in F ratios constructed for other factors.

Examination of the averages indicated that the difference was the result of a consistent tendency to overshoot (stop the movable pipe section after moving past the position of the stationary pipe section) when starting with a 50-mm separation, and undershoot (stop the pipe section before it moves to the position of the stationary pipe section) when starting with a 100-mm separation. The initial position of the movable pipe section in relation to the camera (between the camera and the stationary pipe section or on the other side of the stationary pipe section from the camera) had no effect. In light of these results,

Table 2. Results of the preliminary ANOVA for the depth perception experiment

Source	D.F.	SS	MS	F
Within Subjects				
System	3	497.79	165.93	00.40
Position	3	22881.18	7627.06	58.19[a]
Sys X Pos	9	2947.21	327.49	01.80
Between Subjects				
Observer	3	1700.66	566.89	
Obs X Sys	9	3713.97	412.66	
Obs X Pos	9	1179.61	131.07	
Obs X Sys X Pos	27	4898.88	181.44	

[a] $\alpha = 0.01$

the data from trials beginning with a 100-mm separation between pipe sections were analyzed separately from the data from trials beginning with a 50-mm separation between pipe sections.

2.2.1 Data from 50-mm Separation Trials

Table 3 lists the standard deviations and averages from the data collected during trials which started with a 50-mm separation between pipe sections. These data are averages of statistics calculated separately for each operator. The data indicate that there were no important differences between the performance of standard-resolution, monochromatic TV and the two high-definition TV systems. However, there does appear to be a performance advantage for the stereoscopic TV system. The average standard deviation for this system is 5.70 compared to 9.70 for SRMT, 8.10 for HDCL, and 7.48 for HDMO, indicating that operators were able to position the movable pipe section with more accuracy with stereo television than with the other systems.

Table 3. Standard deviations and averages of alignment errors for trials beginning with 50-mm separation between pipe sections, depth perception experiment

TV system	Standard deviation (mm)
HDCL	7.48
HDMO	8.10
SRMT	9.70
SRST	5.70

2.2.2 Data from 100-mm Separation Trials

Table 4 lists the standard deviations and averages from the data collected during trials which started with a 100-mm separation between pipe sections. These data are also averages of statistics calculated separately for each operator. The data from the 50-mm trials indicate that there are not important differences between the statistics for standard-resolution, monochromatic TV and the two high-definition TV systems. However, there seems to have been a performance advantage for stereo TV. The average standard deviation for this system was 9.54 compared to 11.66 for SRMT, 10.96 for HDCL, and 11.82 for HDMO. This is consistent with the finding for trials beginning with a 50-mm separation between pipe sections.

Table 4. Standard deviations and averages of errors for trials beginning with 100-mm separation between pipe sections, depth perception experiment

TV system	Standard deviation (mm)
HDCL	11.82
HDMO	10.96
SRMT	11.66
SRST	9.54

2.3 CONCLUSIONS

The standard deviations of the distribution of alignment errors are measures of operator sensitivity to distance differences in the remote area. The size of the standard deviation is related to the size of the area in which the operator can differentiate between object distances. It is a measure of the operator's zone of uncertainty; within the area described by the standard deviation, the operator is unable to consistently discriminate between different positions. The small this area is, the smaller the distance differences detectable by the operator. The standard deviation may serve as a figure of merit for the TV systems in the area of transmission of depth information. For the task used in this experiment, the best performance was achieved when operators used the stereo TV system. It produced the lowest standard deviation of alignment errors, indicating that operators using the stereo system were able to reliably detect smaller differences in distance than operators using the other TV systems.

While the depth perception task showed a significant performance advantage for stereo TV, it may be that the task emphasized the type of depth information provided by stereo TV. Because the visual field provided by the task was simple and structured, it may be that cues improved by the HDTV system had no effect within the task. Stereo TV provides cues from retinal disparity, which continue to function in such a visual environment. Other cues are derived by comparing perspective, size, texture, and patterns of light and

shadow. These cues may be enhanced by the higher resolution provided by HDTV, but this type of information was present only to a very limited extent in the simple task used in the experiment. The relative performance of the TV systems may differ in more complex environments because of the availability of other depth cues.

3. EXPERIMENT 2: REMOTE HANDLING

In this experiment the TV systems were compared in terms of how well operators using them performed a simple remote handling task.

3.1 EXPERIMENTAL DESIGN

3.1.1 Experimental Task

The task required operators to pick up four 1.75-in. military connectors (with the threaded rings removed) and plug them into four sockets. Two of the sockets were mounted on a horizontal sheet of aluminum, and the other two were mounted on vertical sheets of aluminum. The vertical sheets were welded to the horizontal sheet at 90° angles and were welded to each other at a 90° angle. The sheets formed three sides of a half-cube, with the sockets mounted on the inside of the cube. Figure 2 is a photograph of the task.

ORNL–PHOTO 8484–85

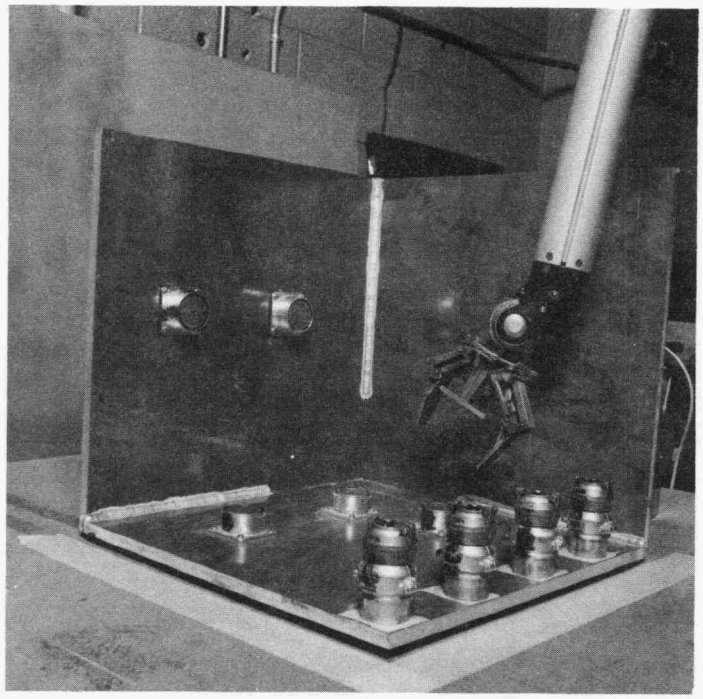

Fig. 2. Remote handling task.

3.1.2 Dependent Variables

Two measurements of performance were made: the time required to complete each task and the number of errors committed during task performance. The latter was recorded by an observer during testing. The observer recorded occurrences of four types of errors: misalignments, misses, drops, and damage. Misalignments occurred when an operator attempted to position the manipulator (or a connector held in the grasp of the manipulator) and failed to achieve the correct position. For example, the operator might attempt to insert a connector into a socket and position the connector off center. Misses included failures to grasp, pick up, or push an object, as in failure to grasp a connector while attempting to pick it up. Drops included incidents when a connector accidentally slipped from the grasp of the manipulator and fell to the surface of the task board. Damage included any events that led to visible damage to the manipulator or connectors.

Task completion time was converted to the logarithm to base 10 of task completion time for analysis, a standard procedure for analysis of duration data because conversion to logarithms prevents outliers (usually high or low scores) from unduly influencing averages and variances.[4]

3.1.3 Operators

Six inexperienced manipulator operators participated in the experiment. Each operator performed the task five times with each TV system, in a set of two 1-h sessions. The order of presentation of TV systems was counterbalanced to prevent the effects of practice from obscuring the effects of the TV systems. The order of presentation for each subject is listed in Table 5.

Table 5. Order of viewing system presentation for the remote handling experiment

	Day 1		Day 2	
Operator	Set 1	Set 2	Set 3	Set 4
1	HDCL	HDMO	SRMT	SRST
2	HDMO	HDCL	SRST	SRMT
3	SRMT	SRST	HDMO	HDCL
4	SRST	SRMT	HDCL	HDMO
5	HDMO	SRMT	SRST	HDCL
6	SRMT	HDMO	HDCL	SRST

3.1.4 Procedures

Operators reported to the Model M-8 manipulator station at the Fuel Recycle Division (FRD) of the Oak Ridge National Laboratory (ORNL). They were met by the experimenter and briefed on the purpose and plan of the experiment. The following instructions were read to them:

This is an experiment to test how well people do remote handling tasks with four television systems. You will do the task you see in front of the manipulators five times with each of the television systems. The first day you will do the task with the _____ television system and then with the _____ system and then the _____ system. While you do the task, I will record the time it takes you to complete the task. Before you start, please do the task once by hand while I time you.

The blanks in the instructions to operators were filled according to the order of administration listed in Table 5.

3.2 RESULTS

An ANOVA[4] was performed on the average of each operator's score in each session. Table 6 presents the results of the ANOVA for the task completion time variable (for the ANOVA, task time was transformed to its logarithm to base 10). Note that one row in Table 6 does not show F tests; this is a factor for which no test exists in the model used.

Table 6. Results of the ANOVA for the logarithm to base 10 of task completion time

Source	D.F.	SS	MS	F
TV	3	0.14	0.05	2.27
Operator	5	0.20	0.04	1.90
TV by operator	15	0.31	0.02	

It represents errors of measurement and is used as a denominator in F ratios constructed for other factors. There were no significant differences between the TV systems on this variable for this task. Figure 3 illustrates the mean difference between systems for the time in seconds required to perform tasks. Figure 3 shows that, on average, operators completed tasks more quickly with HDTV than with the other TV systems; however, this difference was not statistically significant. Figure 4 shows why. Upon examination of individual operator performance, it seems that while some operators responded favorably to HDTV, others performed no better with it than with standard-resolution TV. In spite of the average differences between systems shown in Fig. 3, these data do not consistently support the hypothesis that HDTV leads to faster performance of the remote handling task.

Table 7 presents the results of the ANOVA for the errors variable. Note that one row in Table 7 does not show F tests; this is a factor for which no test exists in the model used. It represents errors of measurement and is used as a denominator in F ratios constructed for other factors. There was a significant difference between the TV systems for this variable; Fig. 5 shows the mean number of errors per trial for each of the systems. Pair comparisons performed on the data indicate that there was no difference between the two HDTV systems (color and monochromatic), but that the two standard-resolution moochromatic systems (monoscopic and stereo) both produced a higher average number of

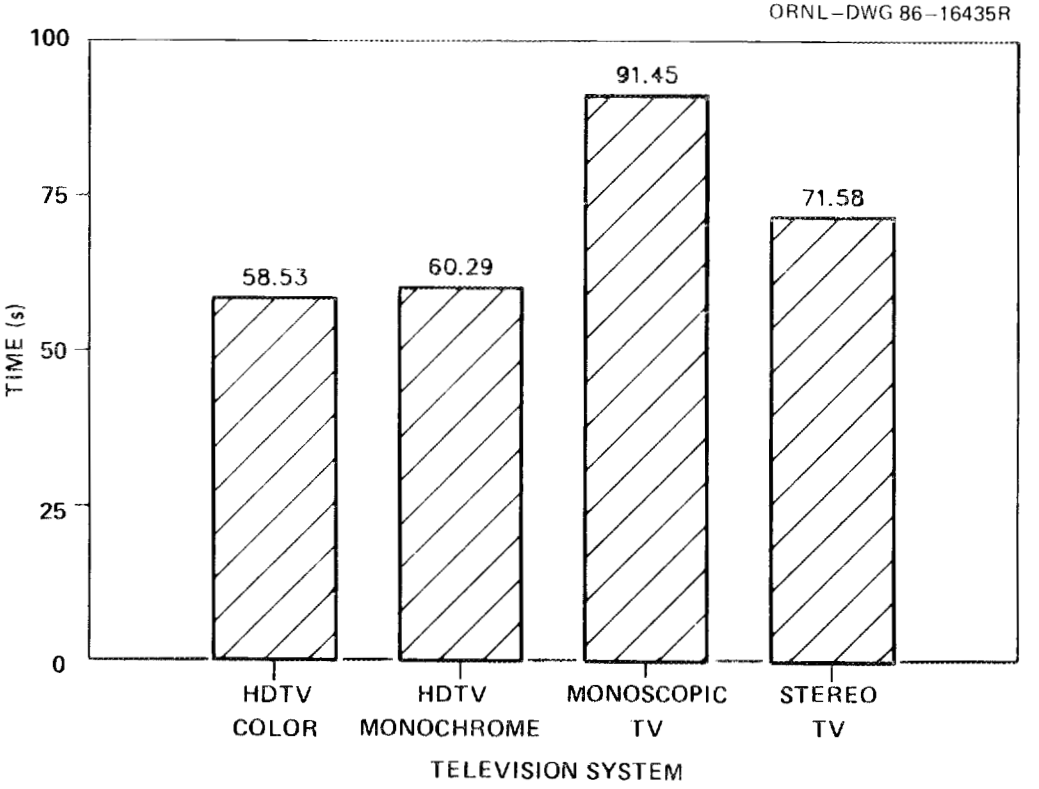

Fig. 3. Mean time to complete the task with each TV system.

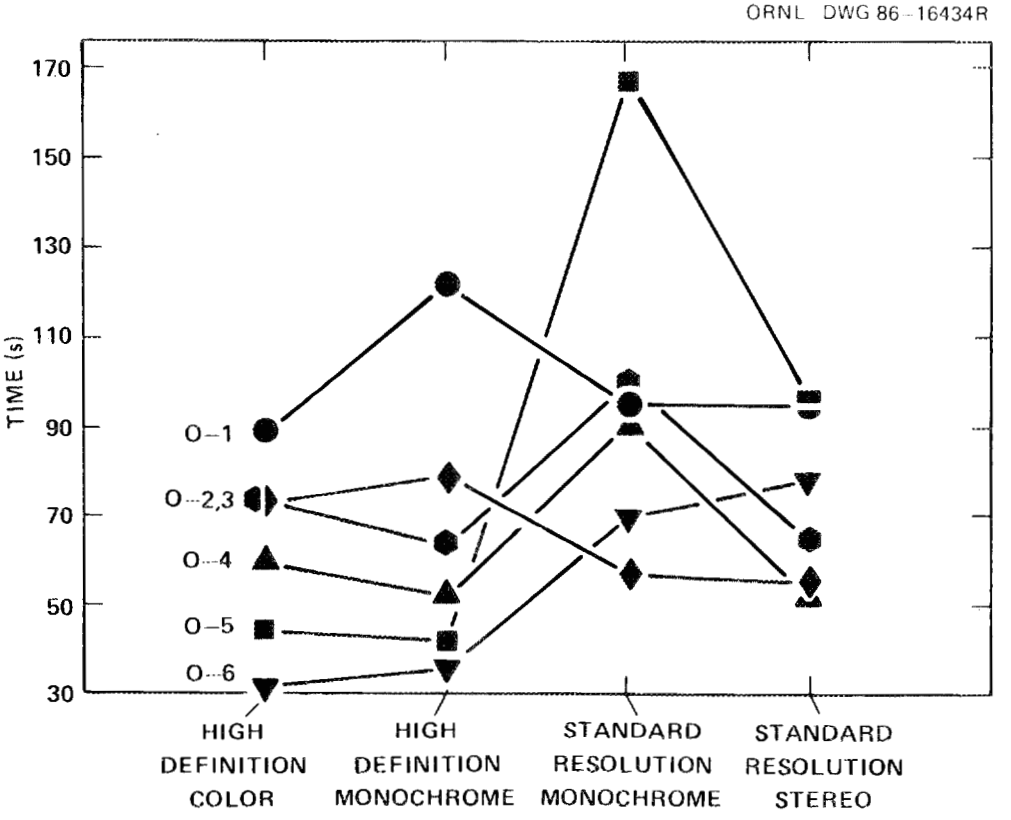

Fig. 4. Mean time to complete the task for each operator and TV system.

Table 7. Results of the ANOVA for the mean number of errors per trial

Source	D.F.	SS	MS	F
TV	3	78.87	26.29	8.72[a]
Operator	5	41.94	8.39	2.78
TV by operator	5	45.22	3.01	

[a] $\alpha < 0.01$

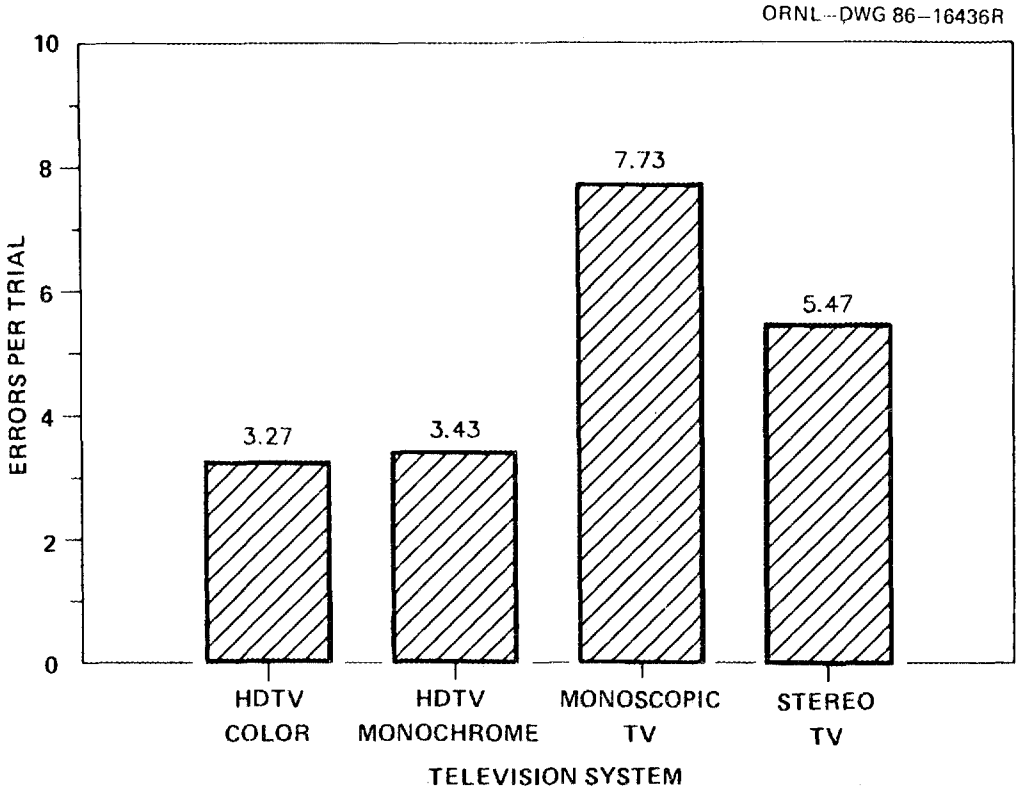

Fig. 5. Mean number of errors per trial for each TV system.

errors per trial. Operators using the monoscopic system also committed more errors than when using the stereo system.

3.3 CONCLUSIONS

High-definition TV allowed the operators to work more accurately than with the other systems, with no sacrifice of task completion time. The sharper edges provided by the ultra-high-resolution HDTV systems allowed the operators to make more accurate judgments of the relative positions of the task components.

4. EXPERIMENT 3: VISUAL INSPECTION

This experiment compared the accuracy with which operators could discriminate flawed from unflawed glass sample bottles with each of the TV systems.

4.1 EXPERIMENTAL DESIGN

4.1.1 Television Systems

The experiment included the same four TV systems as the first two experiments. Each operator completed one testing session with each TV system. Presentation of TV systems was counterbalanced so that any practice effects that may have occurred did not contaminate the experimental data. Presentation of TV systems followed the schedule given in Table 8.

Table 8. Order of presentation of TV systems, inspection experiment

Observer	Testing session 1			
1	HDCL	SRST	HDMO	SRMT
2	SRST	HDCL	SRMT	HDMO
3	SRMT	HDMO	SRST	HDCL

4.1.2 Experimental Task

Observers examined one of a set of two sample bottles for a flaw of known location and type. An experimenter placed a bottle on the horizontal member of the apparatus used in the remote handling experiment (this provided a realistic background for the inspection task), and the operator grasped and rotated the bottle with the manipulator arm. The operator responded to each bottle by indicating whether there was a flaw present. A testing session comprised 100 trials, with each bottle presented 50 times. The order of presentation of flawed- and normal-bottle trials was presented at random independently for each session. Figure 6 is a photograph of the flawed and normal sample bottles used in the experiment.

ORNL–PHOTO 9628–85

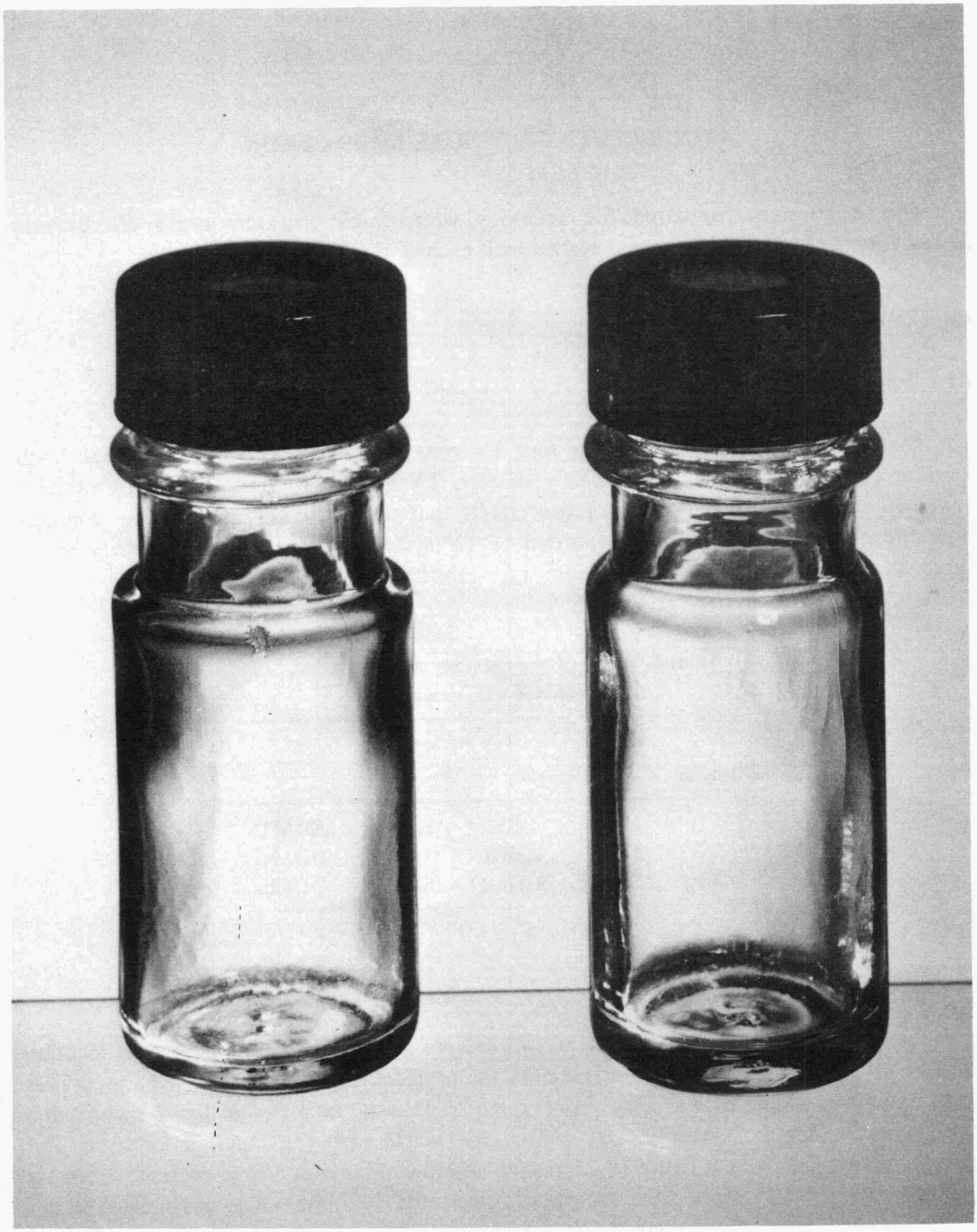

Fig. 6. Sample bottles (flawed bottle on the left).

4.1.3 Participants

Three of the operators who participated in the second experiment participated in this experiment.

4.1.4 Procedure

At the start of each trial in each session, the experimenter indicated to the subject that a trial was about to begin; then he placed a sample bottle in front of the manipulator end-effector. The operator rotated the bottle at will until he or she could make a decision regarding the status of the bottle. The operator indicated his or her decision to the experimenter by saying "flaw" or "no flaw," and the experimenter recorded the decision. At this point, the trial ended, and a new sequence started.

Throughout testing, lighting was constant at the operator station. Lighting in the remote area was adjusted so that each camera operated with about the same aperture size. The size of the image of the bottle on the TV screen was the same on all the television systems. The image of the sample bottle was 1.25 in. (3.175 cm) high, and the image of the flaw was 0.0625 in. (0.1855 cm) high. The operators were seated 36 in. (91.44 cm) from the TV screens. Because the operators were free to move during trials, the exact sizes of the images relative to the visual field are not known. However, at the distance the operators were seated from the screen (36 in.), the sample bottle subtended 3.98° of visual arc when displayed on the TV screens, and the flaw in the bottle subtended 0.099° (0°, 5 min. 58 s) of arc.

4.2 ANALYSIS OF DATA

Analysis of data was by the methods of the Theory of Signal Detection[5] (TSD). The TSD approach yields measures of the sensitivity of operators to objects and of the decision strategy used by operators. One measure of sensitivity involves plotting the probability of hits (correct identifications of flawed items) versus the probability of false alarms (classifying an item as flawed whn it is not). This plot is called an isosensitivity function or receiver operating characteristic (ROC) curve. These ROC curves allow quick comparison of the performance of the television systems. Observer sensitivity can be compared among television systems to determine whether operators are uniformly more accurate with one than with the others; if one outperforms the other systems, it will lead to higher operator sensitivity.

Figure 7 shows the ROC curves for the average of the three operators. The larger the area under the curve, the more sensitive operators were while using the TV system that contributed the data for the curve. Judging from the ROC curve in Fig. 7, the HDTV systems led to greater sensitivity to flaws in the sample bottle than did the stereo and standard-resolution, monochromatic systems. The HDTV monochromatic was slightly better than HDTV with color, but this difference is small. The difference between the systems is not large enough to indicate a real performance advantage for HDTV without color over HDTV with color. Operator sensitivity was slightly higher with standard-

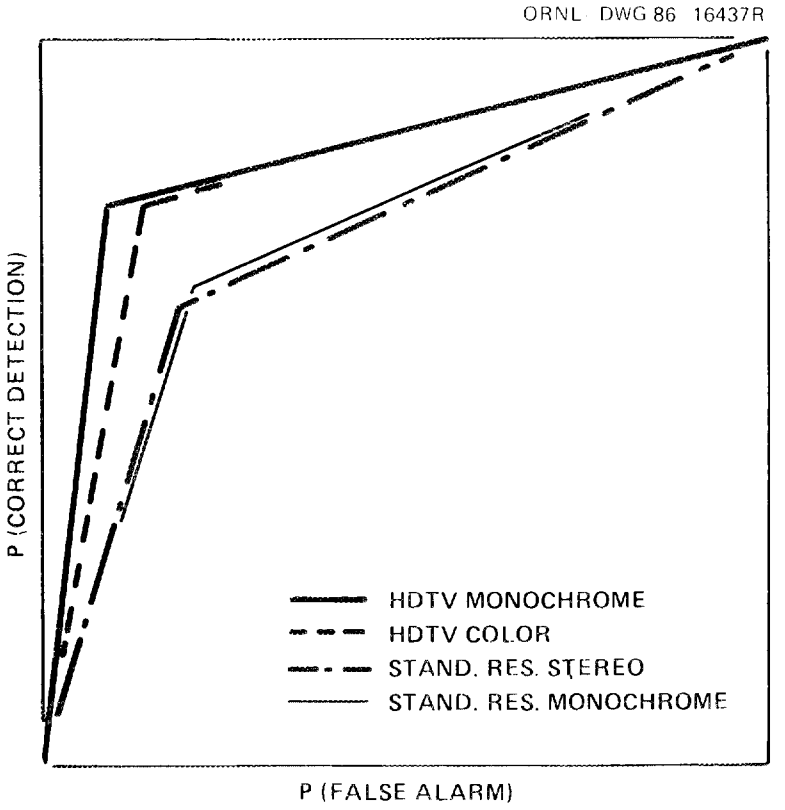

Fig. 7. ROC curve for each television system.

resolution monoscopic television than with the stereo system, but the difference is not large enough to be considered evidence of an advantage for the monoscopic system.

Using TSD, it is also possible to calculate a measure of operator sensitivity. The index of sensitivity is called d' (d prime). The higher d' observed in a set of data, the more sensitive the operator contributing the data. Table 9 lists the average d' calculated for each of the TV systems. Observers using the HDTV systems (HDCL and HDMO) achieved higher d' scores than they did with the other systems, which is consistent with the isosensitivity plots. For HDCL, the average d' was 1.85, and for HDMO d' was 2.05. For SRMT, d' was 1.20, and for SRST d' was 1.18. The performance of the HDTV systems was close,

Table 9. Average inspection accuracy (d') and response tendency (β) for each television system

Television system	d'	β
SRMT	1.20	1.25
SRST	1.18	1.38
HDMO	2.05	1.80
HDCL	1.85	1.33

with a slight advantage displayed by the monochromatic system. The difference between the stereo and the standard-resolution, monoscopic TV systems was very small.

It is also possible to calculate a measure of operator response strategy with TSD. The sensitivity of operators is affected by the decision strategy they use. For example, an operator may attempt to avoid all false alarms; when in doubt, this operator would always respond by indicating that there was no flaw (in this experiment). Another operator might attempt to maximize his sensitivity to flaws; this operator would always respond by indicating that there was a flaw when in doubt. The first operator would have a lower probability of correctly identifying the presence of flaws because this operator always responds "no flaw" when encountering a difficult trial. The second operator would have a high rate of false alarms. Because sensitivity is calculated from data concerning the probability false alarms and the probability of hits, these strategies affect measures of sensitivity. Therefore, it is necessary to examine the operators' response tendencies to ensure that observed differences in ROC curves and d' are not caused by changing operator strategies.

The measure of response strategy is called β (beta). The higher β is, the more the operator was committed to detecting every incidence of a flaw, at the expense of committing more false alarms. Table 9 also lists β for each of the television systems. The figures indicate that on average, β was stable across the television systems. The β for HDMO was somewhat higher than for the other systems, but the difference does not seem large enough to affect the ROC curves or d' calculations.

The β figures can also be used to evaluate the operators' response strategies. In this case, the most efficient response strategy would produce a β of 1.00 because there was an equal probability that a trial might contain a flawed bottle or a normal bottle. The average values of β for this set of operators are all greater than 1.00, indicating that the operators were slightly more concerned with avoiding false alarms than with identifying all instances of flawed bottles.

5. DISCUSSION

The results of the three experiments show that in some areas, HDTV (color or monochrome) can lead to improvement in teleoperator performance. The HDTV is superior to monoscopic, monochromatic, standard-resolution TV and to stereoscopic TV for remote inspection tasks; it is less proficient than stereo TV for interpretation of object distance in the remote area. The HDTV leads to a lower rate of errors committed while performing remote handling tasks, but does not reduce the amount of time necessary to perform the tasks.

These experiments provide an evaluation of HDTV which is fairly wide in scope. The tasks selected are representative of a variety of the tasks that must be routinely performed to remotely maintain a nuclear fuel reprocessing facility. The results and conclusions are applicable in a wide range of situations. However, in attempting to cover so many different situations, the depth with which each was covered was limited. For example, the experiment concerned with visual inspection was limited to one task. In an operating nuclear facility, many inspection tasks would have to be performed, some with very different types of objects to be inspected and very different types of flaws than those used in the experiment. The difference in tasks might produce different results from the pattern reported here. Inspection for cracks in welds often uses penetrant dyes to highlight cracks, for example. In this situation, there might be a performance advantage for color TV systems. Some comments concerning the generalizability of the results from the depth-perception task have already been made. Other tasks that require depth information for successful performance should be defined, and the performance of the different TV systems should be evaluated for those tasks.

Future testing should verify the remote handling task results with different manipulator systems. The CRL Model M-8 used in this testing is quite compliant; if an operator attempted to insert a connector off center, the M-8 was compliant enough to adjust its position to follow the path of least resistance and allow the connector to slip into the socket. Stiffer systems (e.g., the Meidensha BILARM 83A) may require more precise visual alignment, which may increase the performance advantage found for the HDTV systems in this testing.

These experiments failed to address fatigue, an issue which may be an important factor in the performance of TV systems used in remote handling. Future testing should evaluate the relative impact of each of these TV systems in terms of operator fatigue. This testing may be done by examining the decrement in performance from start to end of a task that requires a relatively long time to complete (30 min or more); the difference between initial and final performance is a measure of fatigue. Fatigue may also be measured by evaluating eye strain at the end of long (30 to 60 min) testing sessions. At the end

of such sessions, operators could be administered standard visual acuity tests. Differences in eye strain will be reflected in scores on these tests.

These experiments have outlined a conceptual framework for understanding how HDTV (or any other RV system) performs in remote handling applications. Further testing should concentrate on increasing the depth of understanding of how HDTV performs in the remote maintenance venue.

REFERENCES

1. C. D. Wickens, *Engineering Psychology and Human Performance*, Charles E. Merrill Publishing Company, Columbus, Ohio, 1984.

2. R. A. Erickson, "Line Criteria in Target Acquisition with Television," *Human Factors,* **20**, 573 (1978).

3. J. P. Guilford, *Psychometric Methods*, McGraw-Hill, New York, 1954.

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

5. J. C. Baird and E. Noma, *Fundamentals of Scaling and Psychophysics,* John Wiley and Sons, New York, 1978.

ORNL/TM-10303
Dist. Category UC-86

INTERNAL DISTRIBUTION

1. J. O. Blomeke
2. D. A. Bostick
3. E. C. Bradley
4–6. W. D. Burch
7. G. H. Burger
8. T. W. Burgess
9. D. A. Costanzo
10. S. F. DeMuth
11–15. J. V. Draper
16. M. J. Feldman
17–18. Y. Fujita
19. R. W. Glass
20. N. R. Grant
21. W. S. Groenier
22. W. R. Hamel
23. D. C. Hampson
24. J. H. Hannah
25. T. L. Hebble
26–30. J. N. Herndon
31. W. F. Johnson
32. L. J. King
33. E. H. Krieg, Jr.
34. S. A. Meacham
35. D. R. Moser
36. F. L. Peishel
37. J. C. Rowe
38–39. R. W. Sharpe
40. B. H. Singletary
41. G. E. Smith
42. J. L. Snyder
43. J. G. Stradley
44–45. T. J. Williams
46. H. R. Yook
47–48. Laboratory Records
49. Laboratory Records, ORNL-RC
50. ORNL Patent Section
51. Nuclear Safety Information Center
52. Central Research Library
53. ORNL–Y-12 Technical Library Document Reference Section

EXTERNAL DISTRIBUTION

54. 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
55. F. P. Baranowski, 1110 Dapple Grey Court, Great Falls, VA 22066
56. A. Bejczy, Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109
57. J. Byrd, E. I. du Pont de Nemours and Company, Savannah River Laboratory, Aiken, SC 29801
58. M. M. Clarke, Rockwell International, 12214 Lakewood Blvd., Downey, CA 90241
59. Y. Fujita, PNC Tokai Works, 4-33 Muramatsu Tokai-Mura, Naka-gun, Ibaraki-Ken Japan
60. M. Maeda, PNC Tokai Works, 4-33 Muramatsu Tokai-Mura, Naka-gun, Ibaraki-Ken Japan
61. H. L. Martin, Telerobotics International Inc., 8410 Oak Ridge Highway, Knoxville, TN 37931
62. H. McCain, NASA, Goddard Space Flight Center, Greenbelt, MD 20771
63. A. J. Meintell, NASA, Langley Research Center, Hampton, VI 23665
64. J. L. Nevins, Draper Laboratories, Massachusetts Institute of Technology, Cambridge, MA 02139

65. M. J. Ohanian, Associate Dean for Research, College of Engineering, 300 Weil Hall, University of Florida, Gainesville, FL 32611
66. J. E. Pennington, NASA, Langley Research Center, Hampton, VI 23665
67. R. Pepper, Naval Ocean Systems Center, Ocean Engineering Division, P.O. Box 997, Kailua, HI 96734
68. J. F. Proctor, Senior Technical Specialist, E. I. du Pont de Nemours and Company, Savannah River Laboratory, Aiken, SC 29801

69–80. 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

81. K. Wilshire, NASA, Langley Research Center, Hampton, VI 23665
82. Office of Assistant Manager for Energy Research and Development, DOE-ORO, Oak Ridge, TN 37831

83–175. Given distribution as shown in TIC-4500 under UC-86, Consolidated Fuel Reprocessing Category