Human Factors and Ergonomics in Manufacturing, Vol. 10 (4) 409–430 (2000)

Design of Automation for Telerobots and the Effect on Performance, Operator Situation Awareness, and Subjective Workload

David B. Kaber
Department of Industrial Engineering, North Carolina State University
328 Riddick Labs, Raleigh, NC 27645-7906

Emrah Onal
Lucent Technologies, 1k-525, 101 Crawfords Corner Rd.
P.O. Box 3030, Holmdel, NJ 07733–3030

Mica R. Endsley
SA Technologies, 4731 East Forest Peak, Marietta, GA 30066

ABSTRACT

In this article we review and assess human-centered level of automation (LOA), an alternate approach to traditional, technology-centered design of automation in dynamic-control systems. The objective of human-controlled LOA is to improve human-machine performance by taking into account both operator and technological capabilities. Automation literature has shown that traditional automation can lead to problems in operator situation awareness (SA) due to the out-of-the (control) loop performance problem, which may lead to a negative impact on overall systems performance. Herein we address a standing paucity of research into LOA to deal with these problems. Various schemes of generic control system function allocations were developed to establish a LOA taxonomy. The functions allocated to a human operator, a computer, or both, included monitoring system variables, generating process plans, selecting an "optimal" plan and implementing the plan. Five different function allocation schemes, or LOAs, were empirically investigated as to their usefulness for enhancing telerobot system performance and operator SA, as well as reducing workload. Human participants participated in experimental trials involving a high fidelity, interactive simulation of a telerobot performing nuclear materials handling at the various LOAs. Automation failures were attributed to various simulated system deficiencies necessitating operator detection and correction to return to functioning at an automated mode. Operator performance at each LOA, and during the failure periods, was evaluated. Operator SA was measured using the Situation Awareness Global Assessment Technique, and perceived workload was measured using the NASA-Task Load Index. Results demonstrated improvements in human-machine system performance at higher LOAs (levels involving greater computer control of system functions) along with lower operator subjective workload. However, under the same conditions, operator SA was reduced for certain types of system problems and reaction time to, and performance during, automation failures was substantially lower. Performance during automation failure was best when participants had been functioning at lower, intermediate LOAs (levels involving greater human control of system functions). © 2000 John Wiley & Sons, Inc.

1. INTRODUCTION

Recent literature (Draper, 1995; Endsley & Kaber, 1999; Endsley & Kiris, 1995; Milgram, Rastogi, & Gordski, 1995) concerning research into advances in automation for complex, dynamic systems in general, and telerobots (remote controlled robotic manipulators) specifically, has presented a number of taxonomies of human-centered levels of automation (LOAs). These taxonomies detail control function allocation schemes that may improve systems performance over that resulting from traditional automation. Traditional automation is considered to be the implementation of technology based on its capabilities, but lacking in consideration of the effects of application on a human operator. This approach has been previously justified by the need to reduce human operator workload in many areas including controlling complex manufacturing systems, production planning, semiautonomous materials handling, and so forth. It has also been motivated by the ever-continuing need to increase productivity. Unfortunately, automation researchers have realized over the last three decades that traditional automation has many negative performance and safety consequences associated with it stemming from the human out-of-the-loop (OOTL) performance problem (see Endsley & Kiris, 1995). They have also recognized human-centered LOA as one approach to dealing with this problem.

For example, when traditionally applying automation to a telerobot, responsibility for as many system functions as possible is delegated to a computer and whatever is leftover is given to the human operator. The “left-over” is usually the function of systems monitoring, a task for which it is difficult to develop artificial intelligence or expert systems to accomplish. But this is a function that humans are generally ill suited to perform (Endsley, 1995a) due to extreme susceptibility to vigilance decrements and waning attention. This traditional allocation of functions to human and computer has been associated with problems in operator effectiveness in overseeing automated system functioning and intervening in system operations by taking control from a computer during failure modes (Billings, 1991; Wickens, 1992). Specifically, operators may not detect critical system errors leading to automation failures; they may be inefficient in their responses to failures; they may lack awareness of system states and, consequently, knowledge of how to restore automated functioning due to absence from the direct control loop for extended periods of time. These performance problems not only affect productivity, but they can produce safety concerns as well. By retaining both the human and the computer in system control loops to perform functions for which each server is well suited, operator situation awareness (SA) may be maintained, enabling them to address potential failures, and system performance may be enhanced through computer processing.

1.1. OOTL Performance Factors

OOTL performance problems have been attributed to a number of underlying factors, including human vigilance decrements (Wiener, 1988), complacency (Parasuraman, Molloy, & Singh, 1993; Wiener, 1988), and loss of operator SA (Carmody & Gluckam, 1993; Endsley, 1987; Endsley & Kiris, 1995). Cognitive engineering literature has discussed at length the origins of vigilance decrements (e.g., low signal rates, lack of operator sensitivity to signals) and complacency (e.g., overtrust in highly reliable computer control) in automated systems supervision and has established associations between these human information processing shortcomings and performance problems. When operators are OOTL, vigilance decrements and complacency may both contribute to problems with operator SA. As well, the use of more passive rather than active processing and differences in the type of feedback provided with automated systems may negatively affect operator SA (Endsley & Kiris, 1995).

1.1.1. SA. SA has been defined by Endsley (1988) as, "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and projection of their status in the near future." She has established three levels at which human SA occurs including Level 1 SA—Perception, Level 2 SA—Comprehension, and Level 3 SA—Projection. Achieving Level 1 SA involves perception of information within the environment by allocating attention to sensory cues including task information and distracters. Level 2 SA is attained by relating perceived information to operational goals. Level 3 SA is demonstrated by operator ability to make projections on the state of the environment, as to its figurative position in the future (relative to operational goals). Research (Carmody & Gluckman, 1993; Endsley & Kiris, 1995) has demonstrated OOTL performance to be associated with significant reductions in Level 2 SA in working with automated systems. Carmody and Gluckman (1993) conducted a study in the context of a simulated task battery involving participants simultaneously performing systems monitoring, resource (fuel) management, and compensatory tracking tasks. During the experiment, either the systems monitoring or fuel management tasks were fully automated. They observed operator Level 2 SA to be substantially lower following task automation, as compared to manual control of all three tasks. Endsley and Kiris (1995) conducted a study in which participants were required to respond to automation failures in a simulated automobile navigation task involving expert system assistance. They found that participants functioning under fully automated and semiautonomous conditions prior to automation failures, participant understanding of the state of the system during failure mode performance significantly degraded, as compared to performing under manual control prior to a failure. In other empirical investigations of the effect on SA of monitoring of automated systems, meaningful losses in SA have also been at Level 2 (Endsley & Kaber, 1999).

This and other research has used query-based techniques to obtain objective measures of SA. The queries are directed at the three levels of SA. Operator awareness is evaluated by comparing their responses to questions with factual data on the system. Endsley (1988) formalized a query technique for SA assessment named the Situation Awareness Global Assessment Technique (SAGAT). The SAGAT has been demonstrated to have empirical validity across many different simulated task domains, including aircraft piloting and air traffic control (see Endsley, 1995b, for examples). By determining the information requirements of an operator in relation to task goals, specific queries can be constructed to measure SA on select parameters of a system or global SA developed across all elements of a task. In this way SAGAT measures the construct it claims to measure and is not a reflection of other processes (Endsley, 1995b).

Endsley (1995b) conducted two experiments using SAGAT to examine fighter pilot SA with different cockpit display types. She found the process of querying participants through simulation freezes did not substantially alter SA for subsequent task performance. She demonstrated no ill effects of SAGAT on participant performance with as many as three freezes during a 15-min performance period. Further, she found that freezes of between 5 and 6 min in duration did not cause memory decay or intrude on task performance.

1.2. Taxonomies of Level of Automation

The taxonomies of human-centered LOAs presented in automation literature (see Draper, 1995; Endsley, 1987; Endsley & Kaber, 1999; Endsley & Kiris, 1995; Milgram et al., 1995; Sheridan, 1992) have not only been identified as vehicles by which to improve overall system performance, but they have been aimed at addressing OOTL performance and safety problems along with their underlying causes, as well. Most of the taxonomies offer intermediary LOAs falling somewhere between manual control and full automation. These levels are intended to maintain both human and computer involvement in active systems control for improving operator SA and increasing system performance (through computer data processing). The taxonomies present levels by identifying or describing the roles that the human operator and computer are to play in controlling a system.

Sheridan and Verplanck (see Sheridan, 1992) developed a taxonomy of LOAs for human-computer decision making in the context of undersea teleoperation systems. They identified six functions that either a human operator or computer could maintain in controlling a teleoperator, including “gets,” “selects,” “starts,” “requests,” “approves,” and “tells.” These functions were distributed across human and computer in various ways to form 10 LOAs, including, for example, “human does the whole job up to the point of turning it over to the computer to implement,” and, “computer does the whole job if it decides it should be done, and if so, tells human, if it decides the human should be told” (see Moray, 1986).

Sheridan and Verplanck’s (Sheridan, 1992) taxonomy is one of the most descriptive taxonomies found in the literature in terms of identification of “what” the human and computer are to do under the different LOAs and “how” they work together. However, none of the levels presented in this list has been empirically assessed as to its influence on teleoperator performance or human operator SA (in relation to failure mode performance) in an attempt to differentiate one level from any other.

Endsley (1987) presented a taxonomy of LOA developed in the context of the use of expert systems to supplement human decision making for automated systems control. She identified five functions either a human operator or expert system could play including “suggest,” “concur,” “veto,” “decide,” and “act.” She offered five LOAs by structuring allocation of these roles to both servers ranging from “Manually”—human decides and acts with no assistance from the system, to “Full Automation”—the expert system decides and acts with no operator interaction. Intermediary levels included “Decision Support”—the human decides and acts under suggestions by the expert system, “Consensual AI”—the expert system decides and acts with the concurrence of the operator, and “Monitored AI”—the expert system decides and acts unless the human exercises a veto.

Endsley and Kiris (1995) empirically assessed the effect of the LOAs in this taxonomy on the OOTL performance problem and SA in a simulated automobile navigation task. Their objective was to identify LOAs that sufficiently maintained human operators in the control loop during normal system functioning to permit manual task performance during automation failures. They found performance problems to be more significant under fully automated conditions than under intermediate LOAs. They also found that using lower LOAs, which maintained human operator involvement in active control, was beneficial to SA and participants were better able to perform tasks manually when needed.

Endsley and Kaber (1999) presented a taxonomy of LOAs developed by allocating to either a human, or a computer, or both, generic control functions including “monitoring,” “generating,” “selecting,” and “implementing” based on the capabilities of each server to perform the functions. These functions were identified for use in developing LOAs by studying an array of dynamic-control tasks including aircraft piloting, teleoperation, complex manufacturing systems control, and process control. In addition to identifying these common functions across the named task domains, common operational characteristics were identified to ensure the functions had similar relative importance to, and frequency of use in, systems control. These characteristics included:

1. high subtask demands under limited time resources,
2. operator’s having multiple goals to be pursued simultaneously, and
3. multiple tasks competing for an operator’s attention with each having different relevance to goals.

Endsley and Kaber (1999) formulated 10 LOAs feasible for use in the context of teleoperations. These LOAs make up the taxonomy levels and are presented in Table 1. They have been empirically assessed as to their effect on human-machine system performance, and operator SA and workload, in a dynamic control task. They have also been studied as to their potential for facilitating a smooth transition (in terms of performance) between normal operations and simulated automation failures. Endsley and Kaber (1999) found human-machine system performance to be enhanced by automation that provided computer aiding in the implementation aspect of the task or allocated the implementation role to the computer. With respect to performance during failure modes, the authors found (opposite to the results obtained under normal operating conditions) that human control was significantly superior when preceded by functioning at LOAs involving the operator in the implementation aspect of the task, as compared to being preceded by higher LOAs. Improved SA and lower levels of overall task demand corresponded with higher LOAs.

Draper (1995) presented a taxonomy of levels of control (automation) combining human operators with machine control in a teleoperator capable of both semiautonomous and robotic (fully automated) functions. He identified five different teleoperator functions, including “programming,” “teaching,” “controlling,” “commanding,” and “monitoring,” to be carried out by the human, and four functions to be allocated to the machine, including “controlling,” “modifying,” “communicating,” and “displaying.” The assignment of these functions to the servers was largely dependent upon the capabilities of the technology. Draper (1995) identified five LOAs ranging from total human control to “strategic” control involving human long-term operations planning accompanied by machine performance of tasks. Intermediary levels included:

1. “Manual Control with Intelligent Assistance”—human control and teaching with machine modification of control inputs;
2. “Shared Control”—human control and monitoring and machine control of (routine) subtasks; and
3. “Traded Control”—this level involves consecutive assignment of subtask control to the human and machine depending on the characteristics of the task and server capability.

These LOAs have not been empirically assessed as to their effect on human operator–teleoperator performance or operator SA.

Another taxonomy of levels of autonomy was developed by Milgram et al. (1995) in the context of telerobot control. They structured five LOAs by considering the different roles a human operator could play in telerobot control, including decision maker and direct controller. Their levels ranged from “Manual Teleoperation” to “Autonomous Robotics” including intermediate levels of “Telepresence,” “Director/Agent Control,” and “Supervisory Control.” Milgram et al. (1995) relied on Sheridan’s (1992) definition of “Supervisory Control” for their taxonomy. “Director/Agent Control” was considered to be a form of “Supervisory Control” involving the human operator acting as a director of task performance and the telerobot serving as the agent. The description of “Director/Agent Control” does not specifically state what the human or computer are to do. The

TABLE 1. Endsley and Kaber's (1999) LOA Taxonomy

	Functions			
Level of Automation	Monitoring	Generating	Selecting	Implementing
Manual Control	Human	Human	Human	Human
Action Support	Human/Computer	Human	Human	Human/Computer
Batch Processing	Human/Computer	Human	Human	Computer
Shared Control	Human/Computer	Human/Computer	Human	Human/Computer
Decision Support	Human/Computer	Human/Computer	Human	Computer
Blended Decision-Making	Human/Computer	Human/Computer	Human/Computer	Computer
Rigid System	Human/Computer	Computer	Human	Computer
Automated Decision Making	Human/Computer	Human/Computer	Computer	Computer
Supervisory Control	Human/Computer	Computer	Computer	Computer
Full Automation	Computer	Computer	Computer	Computer

authors considered “Manual Teleoperation” to be any control situation that constrained the human operator to remain continuously in the direct control loop. They seemed to consider this mode of control to be burdensome to the operator and to imply the only reason for its existence is that the current limitations of advanced control technologies prevent more common types of telerobotic operations (e.g., hazardous materials handling, undersea structures maintenance, and space station assembly) from being completely automated. In contrast to their goal, the case has been made here for the use of direct control to maintain operator system awareness for performance during failures and to account for the imperfect reliability of technology. Therefore, even if fully automated performance of teleoperations is available, it may be not be advantageous to overall system performance.

Milgram et al. (1995) also identified “Telepresence” as a LOA at which the human operator is provided the means to conduct teleoperations as if (s)he were actually present at the remote site. Telepresence has been previously described (Draper, Kaber, & Usher, 1998) as a mental state involving the transport of one’s consciousness to a remote site that is influenced by the characteristics of the system a human is working with, the task he or she is performing, and his or her innate abilities and personal experiences; however, its identification as a LOA is novel. Milgram’s et al. (1995) explanation of telepresence does not establish the exact roles the human and machine are to play in teleoperation or telerobotic control. Rather they state that it involves a master-slave control system and the use of a helmet-mounted display for immersion of an operator’s senses in stimuli from the remote environment. This description may be labeled more appropriately as a teleoperator system configuration to motivate telepresence rather than a LOA. Further, even if human and computer responsibility for various teleoperation system functions were established, a LOA motivating telepresence may not necessarily involve the human operator maintaining the same decision-making or control functions all the time. Different function allocation schemes may influence telepresence in different ways, but it seems unlikely that a single LOA represents telepresence for all systems and task circumstances. Milgram et al. (1995) have not empirically assessed “Telepresence” or any of the other LOAs in their taxonomy for their effect on telerobot control performance, or measures of human operator SA or telepresence.

Anderson (1996) indirectly put forth a taxonomy of LOAs for robot systems including “Autonomous Control,” direct “Teleoperation,” and “Shared Control.” In “Autonomous Control” an operator programs a series of points that a robot is to move to and perform manipulative functions. Under pure “Teleoperation” an operator is required to directly command all motions of the robot in real-time using a hand-controller (e.g., SpaceBall®) instead of programming positional goals. “Shared Control” of the robotic system involves a blend of the characteristics of these two modes including superimposing inputs of the operator and computer control on each other. None of the LOAs that Anderson (1996) described have been empirically assessed as to their effect on teleoperator performance.

1.2.1. Discussion. Criticism has been made of several of the above described taxonomies that they may be incomplete in different senses and may have limited applicability to specific types of systems for improving performance and abating the negative consequences of OOTL performance. It has been suggested that taxonomies of LOA be developed for specific tasks to promote their usefulness in resolving real-world automation problems. Such efforts might lead to OOTL performance problem reductions in the task for which a taxonomy is developed; however, they will not lead to a generalized theory, or method of, automation in complex, dynamic systems that may adequately serve designers working in a broad spectrum of automation applications.

Additional criticism has been offered in that the capabilities of computers are continually changing, including their capacity to perform advanced control functions such as scheduling operations and selecting “optimal” schedules from several alternatives (decision making). As these capabilities are enhanced, existing LOA taxonomies may become inefficient and ineffective in structuring human and computer control of systems. However, in using LOA as an approach to automation it is important to recall the need to consider performance capabilities during both normal operations and failure modes. This alone may sustain the usefulness of certain existing LOA taxonomies in future systems.

To date, a limited amount of experimental work has been conducted to investigate the usefulness of intermediary human-centered LOAs for enhancing specific task performance, or to examine the effects of LOAs on operator SA and mental workload. This makes it difficult to expressly quantify, for example, the impact of “Shared Control” (Anderson, 1996; Draper, 1995; Endsley & Kaber, 1999) on human perception and understanding of automated systems information and, consequently, operator manual control performance during an automation failure. As well, the issue of LOA obsolescence due to advances in computer technology has not been addressed in longitudinal studies.

The studies of Endsley and Kiris (1995) and Endsley and Kaber (1999) demonstrate that traditional, full automation of a system or task may not be advantageous if joint human-machine performance is to be optimized. They also offer support to the usefulness of intermediate LOAs presented in general taxonomies of control for specific tasks and functions in order to keep human operators’ SA at higher levels and to allow them to perform critical functions during failures (Endsley, Kaber, & Onal, 1997).

The purpose of the present research is to further examine the benefits of intermediary LOAs, specifically in a high-fidelity simulation supporting generalizability of results to a real-world application. Further, it was intended to demonstrate the usefulness of general LOAs in the context of a specific application. This was accomplished by assessing the impact of LOA on telerobot performance under both normal operating conditions and failure modes, and its effect on operator SA and subjective workload.

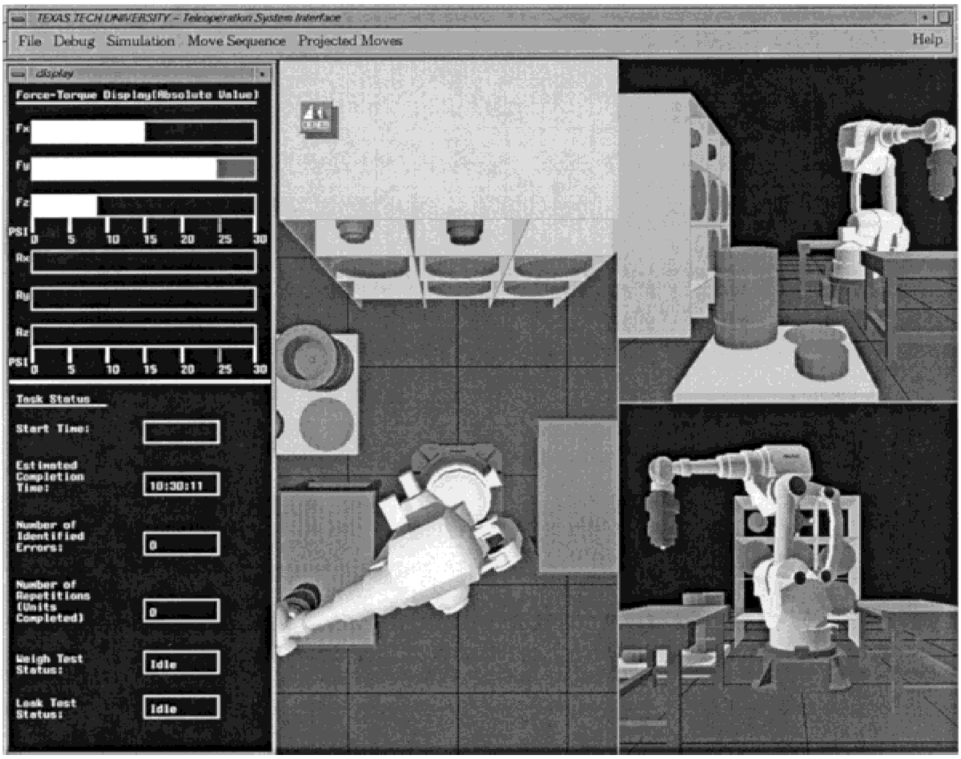
The use of telerobots under intermediate LOAs for nuclear materials handling is particularly appropriate for this research. The overall objective of using a remote manipulator system to perform such a task is to improve operator safety and task performance by reducing human radiation exposure and the potential for certain system errors, such as dropping materials. The telerobot allows the human to work outside the task environment, limiting harmful effects of radiation on the body. From the perspective of this research, intermediate LOAs can be applied to the telerobot to blend human and computer control, thus maintaining both servers in the loop. It was hypothesized that joint human-computer control would serve to maintain operator SA, and efficient and effective system recovery would be made in the event of automation failure. This is critical because there is zero tolerance for errors in this application environment.

2. EXPERIMENT

2.1. Task

An experiment was conducted in which the LOA taxonomy in Table 1 was explored using a high-fidelity simulation of a telerobot performing safety tests on plutonium storage containers. The simulation required participants to interact with a computer in controlling a Fanuc® S-800 robotic arm. Operator control was facilitated through a graphical user interface (GUI) designed by Kaber et al. (1997), shown in Figure 1. In the simulation, an operator controlled the robotic arm in the following tasks:

1. removing storage containers from a staging rack one at a time using a vacuum gripper integrated with the arm,
2. placing each container on a pedestal adjacent to the staging rack and unbolting its lid with a wrench gripper,
3. removing the lid and packing material from the storage container using the vacuum gripper,
4. removing a plutonium containment vessel (CV) from the storage container using a pneumatic gripper,
5. placing the CV at a weigh station and weighing it for 15 s,
6. placing the CV at a leak testing station using the pneumatic gripper and testing it for 10 s,
7. repacking the CV in the storage container using the pneumatic gripper,
8. returning both the packing material and lid to the container using the vacuum gripper and bolting the lid using the wrench, and
9. returning an inspected container to the staging rack.



Start Time:

Estimated Completion Time: 10:30:11

Number of Identified Errors: 0

Number of Repetitions (Units Completed) 0

Weigh Test Status: Idle

Leak Test Status: Idle

Figure 1 Graphical-user interface for telerobot control in simulated nuclear materials handling.

The operators' goal in the simulation was to safely and efficiently handle the plutonium containers through the inspection process. They were to minimize the number of handling errors including collisions between CVs and other task objects, and dropping a CV. Further, they were to minimize the average cycle time for processing a container, thereby maximizing the number of CVs inspected per simulation trial.

2.2. Equipment

The telerobot task was simulated on a Silicon Graphics® (SGI) Indigo^2; workstation using a 21-in graphics monitor operating under 1600 by 1280 resolution. The system was integrated with a mouse, standard keyboard and SpaceBall® controller. The system maintained the simulation with an update rate of 30 frames/s.

An Apple Macintosh, configured with a 15-in. graphics monitor, a standard keyboard and mouse, was used in the study to electronically present SAGAT queries (Endsley, 1988) and NASA-Task Load Index (TLX) demand-factor ranking and rating surveys. The SAGAT and the TLX were intended to establish operator SA and overall perceived workload, respectively, in performing the task.

2.3. Participants

Ten participants (6 men and 4 women) were recruited for this study from the Texas Tech University student population. They participated for monetary compensation on a voluntary basis. All had normal or corrected to 20/20 visual acuity, full-color vision, and computer and mouse experience. They were all also right-hand dominant. (These characteristics were required by the study due to the simulation being a visual presentation of the telerobot and its work environment, and because a right-handed SpaceBall® was used.)

2.4. Experimental Design

This was a repeated measures investigation in which the participants served as replications in examining the effect of LOA as an independent variable in performance during both normal operation of the simulation and simulated failures. As well, participants served as repeated observations of the LOA effect on SA and mental workload. The simulated system was programmed to allow for the use of five LOAs presented in the taxonomy in Table 1 including:

1. “Action Support,”
2. “Batch Processing,”
3. “Decision Support,”
4. “Supervisory Control,” and
5. “Full Automation.”

These levels represented a range of automated system control function allocations from human manual control with computer assistance (Action Support) to complete computer control (Full Automation) (Endsley & Kaber, 1999). The five levels were defined in terms of the functions of the telerobot and the simulated task. Definitions for all levels are presented in Table 2.

TABLE 2. Descriptions of LOAs for Telerobotic Control in Nuclear Materials Handling

LOA	Description
Action Support (AS)	Human generates container-processing plan (e.g., unpack container, weigh CV, leak test CV, etc.). Human (mentally) selects CVs and inspection tasks (e.g., move CV to weigh station) to achieve plan. Human implements plan by using SpaceBall® to control telerobot in moving CVs about workcell and activating computer implementation of specific task functions including picking-up, returning, bolting, and unbolting task objects. All other input devices to telerobot (e.g., keyboard) are disabled and computer implementation of task functions is restricted to object approach distances less than 30 cm. (AS, therefore, provides limited joint human/computer control of the teleoperator in carrying out task implementation.)
Batch Processing	Human generates a container processing plan and selects inspection tasks (telerobot move sequences) to be implemented by computer. Human uses mouse controller and buttons to select move sequence options from pull down menus presented in GUI. Selected move sequences are added to list of scheduled moves allowing for human advanced processing of tasks. Computer implements selected tasks by automatically controlling telerobot. (This LOA, therefore, provides full automation of the implementation portion of the process.)
Decision Support (DS)	Human and computer both generate a container-processing plan. Human develops own plan by using mouse controller to select move sequences from pull down menus. Human decides whether to use computer-generated plan or own plan. Human can select computer plan by using mouse. Computer implements human selected plan by automatically controlling telerobot. (DS, therefore, provides a higher LOA by incorporating human and computer process plan generation with human selection and computer implementation.)
Supervisory Control (SC)	Computer controls all functions with human over-ride capability. Computer generates container processing plan by considering all task variables (e.g., CV present at weigh station, time remaining to weigh CV, etc.). Computer selects robot arm move sequences to facilitate inspection, and implements processing plan by automatically controlling teleoperator. Human can intervene in control process if (s)he thinks computer is not safely, effectively and efficiently processing CVs. Human intervention is accomplished by depress of key on keyboard halting robot processing of current move and shifting LOA to DS. Return to SC is also accomplished via keyboard. (This LOA is, therefore, representative of SC systems in which the control functions are mostly automated, but where human intervention is expected.)
Full Automation (FA)	Computer performs all functions including system monitoring, container processing plan generation, task or move sequence selection, and plan implementation. Human intervention is not permitted (all input from keyboard and hand-controller is ignored). (Therefore, under FA the human can only observe the system.)

2.4.1. Dependent Variables. The dependent variables recorded during the experiment included time-to-container completion and the number of processing errors (robot and task object collisions) under normal operating conditions. These variables were averaged across a single trial (excluding the time spent in failure modes using Action Support to control the robot) for all 50 trials (5 LOAs × 10 participants).

Time-to-system-recovery and the number of actions executed toward recovery were also recorded during automation failures. Observations were made for two failures simulated during 40 of the 50 test trials (4 LOAs [Batch Processing, Decision Support, Supervisory Control, and Full Automation] × 10 participants). Observations were not collected on the 10 trials involving participant control of the robot using Action Support because the LOA for normal operations did not differ from that used during failures.

Operator SA was measured during the study using SAGAT (Endsley, 1988) queries regarding the three levels of SA proposed by Endsley (1988). The queries used during the experiment and possible responses are shown in Table 3. Operator perceptions of the system were evaluated by comparison with factual simulation data recorded using the SGI workstation. Through these comparisons, SAGAT served as an objective measure of SA (Endsley, 1995b). The queries were posed to participants during three freezes across all 50 test trials. SA was quantified as the percentage of correct responses to each query. This percentage was averaged across all queries targeted at a particular level of SA (Level 1 SA, Level 2 SA, and Level 3 SA). These data served as composites of operator perception and comprehension of system information, as well as future system state predictions.

The NASA-TLX was used to subjectively assess the overall workload experienced by operators. Prior to experiment events, participants were required to complete pairwise rankings of six mental and physical demand components (in terms of importance to task performance). During the experiment, they completed ratings of the demand components. This data was used to compute a composite workload score (Hart & Staveland, 1988). Workload was assessed during each SAGAT freeze.

2.5. Procedure

This study was conducted across 5 consecutive days. On the first day, participants were instructed in how to control the simulated Fanuc® S-800 arm under direct teleoperation (Action Support) using the SpaceBall®. They were trained for 45 min without any interruptions (e.g., automation failures). This was followed by an additional 60-min practice period requiring the use of Action Support with three simulation freezes occurring at random points in time to administer SAGAT queries. Participants were informed in advance that freezes would occur, but they were not given knowledge of the number of freezes or the interfreeze-interval times. When a freeze happened, the display screen of the SGI workstation was blanked and participants responded to the queries electronically using the Macintosh®. Each freeze lasted until participants completed the SAGAT queries and the NASA-TLX workload rating. Subsequently, they resumed the telerobot simulation. The first day of the study was concluded with participants being tested on their control of the robot using direct teleoperation. This was considered to be the lowest level of control in the study, as it involved human operator performance of all system functions including direct manual control of the movement of the Fanuc® arm. This level did, however, offer some computer assistance in the implementation role, specifically computer guidance in fine-detailed positioning of the manipulator. During the test period, all performance measures were recorded and three simulation freezes were conducted to capture operator SA. NASA-TLX demand component rating forms were administered during each of the SAGAT freezes. No automation failures were encouraged during this period.

The 2nd through the 5th days of the experiment involved training participants in control of the telerobot at different LOAs (i.e., Batch Processing, Decision Support, Super-

TABLE 3. Queries on SA, Potential Responses, and the Level of SA Targeted by Each Query

SAGAT Query	Possible Responses	Level of SA
What is the current gripper type?	(a) vacuum; (b) wrench; (c) pneumatic; (d) none.	1
What is the current (just completed) process?	(a) picking-up drum; (b) picking-up plutonium container; (c) picking-up gripper; (d) picking-up packing material; (e) picking-up lid; (f) returning drum; (g) returning plutonium container; (h) returning gripper; (i) returning packing material; (j) returning lid; (k) moving drum; (l) moving plutonium container; (m) moving gripper; (n) moving packing material; (o) unbolting a drum lid; (p) bolting a drum lid; (q) no current process.	1
Where is the plutonium container that is currently being processed?	(a) weight station; (b) leak check station; (c) between stations/attached to the robot arm; (d) inside drum on pedestal; (e) no plutonium container currently being processed.	1
How many containers have been processed so far (current drum plus completely processed drums)?	(a) 1–2; (b) 3–4; (c) 5–6; (d) 7–8; (e) 9+.	1
Does the last processed plutonium container meet the weight limit standard?	(a) yes; (b) no; (c) no plutonium containers have been processed in this shift.	2
Does the last processed plutonium container exceed the leak check limit standard?	(a) yes; (b) no; (c) no plutonium containers have been processed in this shift.	2
Do you have the appropriate gripper for the current step?	(a) yes; (b) no; (c) no gripper is needed in this step.	2
Is the robot currently close enough to an object to perform a pick-up, return, unbolt or bolt action?	(a) yes; (b) no; (c) none of these actions are needed at this point in the process.	2
What type of gripper do you need to use next (after current gripper)?	(a) vacuum; (b) wrench; (c) pneumatic; (d) none.	3
What process needs to be executed next (ensuring no collisions or errors)?	(a) picking-up drum; (b) picking-up plutonium container; (c) picking-up gripper; (d) picking-up packing material; (e) picking-up lid; (f) returning drum; (g) returning plutonium container; (h) returning gripper; (i) returning packing material; (j) returning lid; (k) moving drum; (l) moving plutonium container; (m) moving gripper; (n) moving packing material; (o) unbolting a drum lid; (p) bolting a drum lid; (q) no current process.	3
How many plutonium containers will you finish processing in this shift?	(a) 1–2; (b) 3–4; (c) 5–6; (d) 7–8; (e) 9+.	3
Where does the plutonium container currently being processed go to next?	(a) weight station; (b) leak check station; (c) drum (on pedestal); (d) rack (inside drum); (e) no plutonium container currently being processed.	3

visory Control, and Full Automation) and testing their performance under normal conditions and failures, as well as assessing their SA and workload in operating the system. During the training, three simulation freezes were used for SAGAT queries and two automation failures occurred at random points in time. Participants were informed in advance that a failure caused by several specific types of system errors might occur; however, no information concerning the number or frequency of failures that could be expected within a trial was provided. When a failure occurred in the simulation, a pop-up message box was displayed on the SGI monitor indicating the type of error that had been encountered including:

1. “robot stuck,”
2. “no electricity,” or
3. “gripper malfunction.”

Participants were required to depress the *r*, *e*, and *g* keys on the keyboard corresponding to the letters of the three error messages. Once the operator responded to an error appropriately, the LOA of the system shifted to direct teleoperation (Action Support). The system remained under this level of functioning until the operator successfully completed an action sequence (e.g., picking up an object with the arm). Subsequent to this, the system resumed operation under the test LOA.

Following training, participants were tested under the same circumstances as training (LOA, number of freezes, failures and errors). In total, each participant devoted approximately 15 hr of time to the experiment.

3. RESULTS AND DISCUSSION

3.1. Performance Under Normal Operating Conditions

At the onset of the experiment, it was hypothesized that intermediate and higher LOAs would produce superior performance in terms of task completion time and collision avoidance due to allocation of the implementation aspect of the task to automation and human decision making in generating processing strategies. An analysis of variance (ANOVA) was conducted on time-to-container completion revealing LOA to be significant in its effect, $F(4,9) = 53.85$, $p < .001$. A plot of the mean time-to-container completion as a function of LOA is shown in Figure 2. In general, performance improved (i.e., task time decreased) as the LOA increased.

Tukey’s honestly significant difference (HSD) tests time demonstrated Action Support to produce significantly shorter times ($p < .05$) than all other levels. Supervisory Control and Full Automation were significantly higher ($p < .05$) than Batch Processing (but not different from each other). This analysis reveals the benefit of computer programmed motion control over the telerobot. Direct teleoperation (Action Support) required human involvement in the implementation aspect of the task (motion path control) and produced the lowest performance. This can be attributed in part to the difficulty participants had in controlling the telerobot using the SpaceBall®. They were required to mentally map three translations and rotations from the hand-controller to the movement of the simulated robot. It appeared to be cognitively taxing for subjects to keep track of and isolate all six different movements during performance. (Extensive training in the use of the Space-

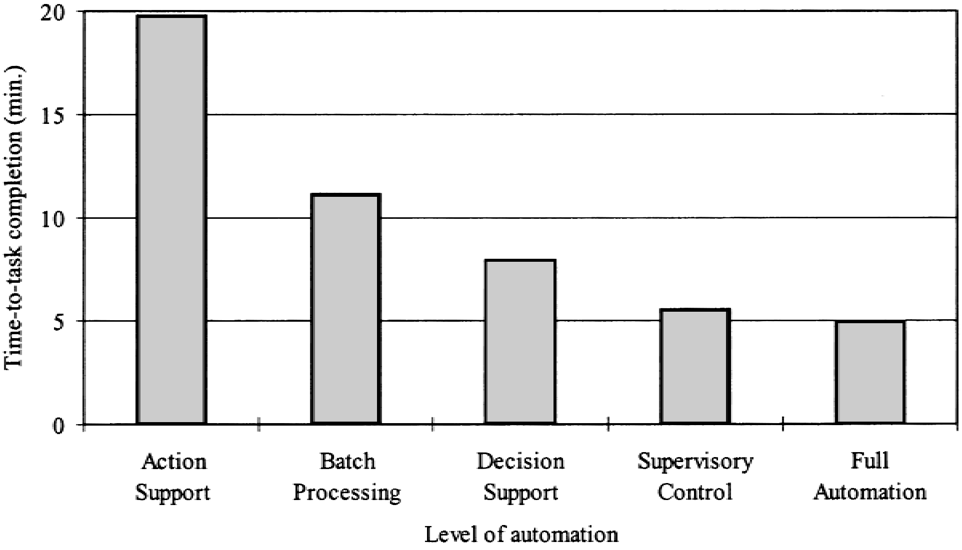


Figure 2 Mean time-to-container completion as a function of LOA.

Ball® for direct teleoperation was provided; however, there did appear to remain an effect of the controller configuration on performance.)

The superiority of Supervisory Control and Full Automation to Batch Processing can be explained in terms of the roles the human and computer maintained during the teleoperation. As presented in Table 2, under Batch Processing, participants were required to generate a container-processing plan and to develop specific move sequences to that plan using a pull-down menuing system in the GUI (see Figure 1). These sequences were posted to a list of moves to be implemented by the computer. Participants also had the capability to clear the list at any time if they detected a potential processing error. Although the system did allow for operator advanced planning of telerobot moves, the participants seldom took advantage of this capability, often waiting for processing of a particular sequence to finish to ensure its success (safety). They tended to adopt a “move and wait” strategy, which is commonly observed in actual teleoperations with control lag. Consequently, Batch Processing never managed to produce performance equivalent to complete computer control under Full Automation or Supervisory Control.

These results suggest that when participants were provided with the control capability to improve task accuracy through an intermediary LOA by preventing incidents such as a collision, they took advantage of it at the expense of processing speed. This is probably a reflection of the instructions given to participants concerning the goals of the task and a priority placed on system safety versus processing efficiency.

An ANOVA on the mean number of collisions during normal functioning of the system also indicated LOA to be significant, $F(4,9) = 5.25$, $p < .01$. A means plot for collisions as a function of LOA is shown in Figure 3. The graph reveals a bathtub trend of the response as driven by LOA with the preponderance of collisions having occurred under Action Support and Full Automation.

Tukey’s HSD test was used to further investigate the effect of LOA on this response revealing Action Support and Full Automation to be significantly different ($p < .05$)

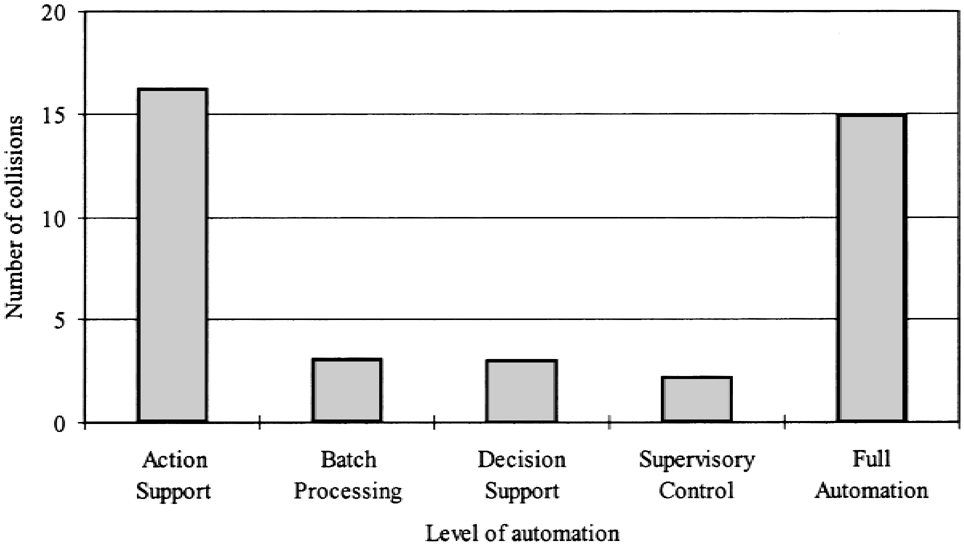


Figure 3 Mean number of robot and task object collisions during normal operations as a function of LOA.

from the intermediate LOAs (i.e., Batch Processing, Decision Support, and Supervisory Control).

The higher number one collisions under Action Support can again be attributed to the complexity of manual control of the simulated robot arm using the SpaceBall®. The multiple comparison results indicate that, when possible, participants capitalized on their control capabilities under Batch Processing, Decision Support, and Supervisory Control to reduce errors. These levels involved the human in process plan generation in various ways combined with the ability to select an “optimal” plan from computer-generated alternatives. The computer implemented the plans. This combination of human decision making with computer processing, in the context of the teleoperation, served to significantly benefit performance accuracy.

The results on the number of collisions demonstrate the usefulness of intermediate LOAs in a system in which automated control may not be perfectly reliable (i.e., the robot could collide with objects). Batch Processing and Decision Support permitted operator cognition on potential error conditions resulting in significant reductions in the number of errors; however, Full Automation resulted in all errors going unchecked. The negative consequence of the latter in the real world is reduced product quality, or, in the case of teleoperator nuclear materials handling, reduced safety.

3.2. Performance during Automation Failure

It was hypothesized that operator efficiency and effectiveness in manually recovering the telerobot from automation failures would be superior when preceded by normal performance under intermediate LOAs maintaining the participant in the control loop and promoting SA. ANOVAs were conducted on data recorded during the simulated automation failure modes in which participants were required to perform direct teleoperation (Action Support) to return the system to a higher LOA. Results revealed a significant effect, $F(3,9) = 3.11$, $p < .05$, of the LOA preceding a failure on the time-to-system recovery from failure. A means plot of the time-to-recovery as a function of LOA is shown in Figure 4. In general, recovery time increased with LOA.

Tukey's HSD test on the LOA effect revealed recovery time subsequent to performance under Full Automation to significantly differ ($p < .05$) from Batch Processing and Supervisory Control, but not Decision Support. These findings can be explained by referring to Table 2. The definition of Supervisory Control involves Full Automation along with human process intervention, when needed, through a shift in the LOA to Decision Support. This makes Supervisory Control equivalent to Decision Support when operators enter the control loop to address perceived errors in computer processing. During the experiment, participants frequently intervened in systems control. Because Decision Support required a greater degree of human involvement in the simulation, these interventions may have increased participant awareness of system states prior to a failure promoting faster recovery times, as compared to Full Automation.

Under normal conditions, Supervisory Control only required the human to monitor the robot while the computer maintained all other functions. It off-loaded system responsibilities including planning and decision making from the human to the computer. These responsibilities included evaluating computer processing plans for safety and efficiency, which may have prevented participants using Decision Support from monitoring system

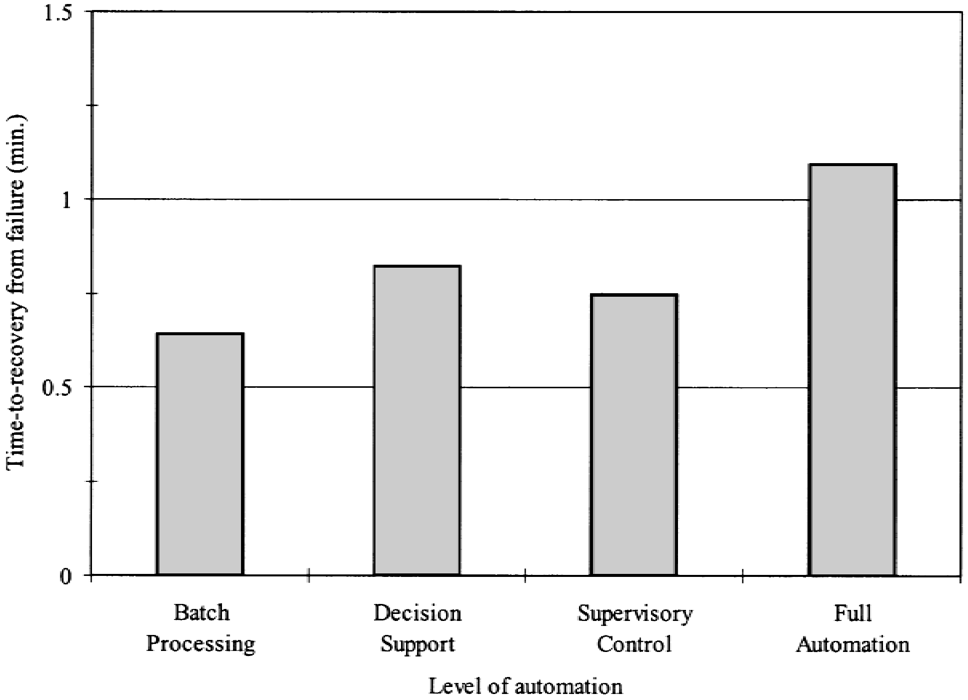


Figure 4 Mean time-to-recovery from telerobot failure dependent upon the LOA used under normal conditions preceding the failure.

states and other task-relevant information and, consequently, relating such information to a failure condition for efficient recovery.

Batch Processing is related to Decision Support in that the human and computer jointly perform monitoring of the system and the computer implements the process plan. However, Batch Processing limits the generating and selecting functions to the human, whereas Decision Support permits both human and computer responsibility for the plan generation aspect of the task. This difference in the function allocation schemes explains the observed difference between the two LOAs in terms of time-to-recovery in the simulation. Batch Processing involved the participants in telerobot control to a greater extent than Decision Support, possibly promoting heightened states of participant system and task awareness for efficient recovery subsequent to a failure.

LOA did not significantly affect the number of control actions executed during a failure to achieve recovery. Across all LOAs, participants were able to identify and execute, on average, a single action (e.g., pick up, return, unbolt or bolt) to cause recovery of the system. The number of actions needed to recover the system was significantly correlated ($r = 0.85$, $p < .05$) with the time-to-recovery.

In summary, these results are supportive of lower level automation (Batch Processing) maintaining human involvement in the control loop during normal system functioning. This promoted operator performance during failure modes in terms of identifying corrective actions to recovery and, consequently, the time-to-recovery. They also suggest that a blending of intermediate (Decision Support) with higher (Supervisory Control) LOAs facilitates control loop involvement and, at the same time, reduces operator workload for perceptual activities, which may be beneficial to failure mode functioning in the context of telerobot control.

3.3. SA

Intermediate LOAs combining human and computer strategizing and control in task planning and execution were anticipated to produce high operator SA as a result of the operator being retained in the telerobot system control loop. ANOVAs were conducted on the average percent correct responses to SAGAT questions covering the three levels of SA. Results revealed a significant effect of LOA, $F(4,9) = 3.4$, $p < .05$, only on Level 3 SA. Figure 5 shows the mean percentage of correct responses to Level 3 SA questions, which significantly decreased from low to intermediate LOAs. According to Tukey's HSD test Supervisory Control and Full Automation were not significantly different from Batch Processing and Decision Support ($p > .05$) in their effect on the percentage of correct responses.

Action Support was characterized by human involvement in all aspects of telerobot control. Action Support varied from Batch Processing in that the latter stripped participants of the capability to manually control the robot using the SpaceBall®. This difference between the two LOAs can be directly associated with a substantial degradation in operator ability to predict the next move sequence in the materials handling task and the type of gripper needed for subtask performance following the move.

Decision Support, like Batch Processing, also did not permit human involvement in implementing planned move sequences, and, as noted previously, it limited their role in the planning function. This was associated with a significant difference ($p < .05$) between Action Support and Decision Support, in terms of Level 3 SA, according to Tukey's HSD test.

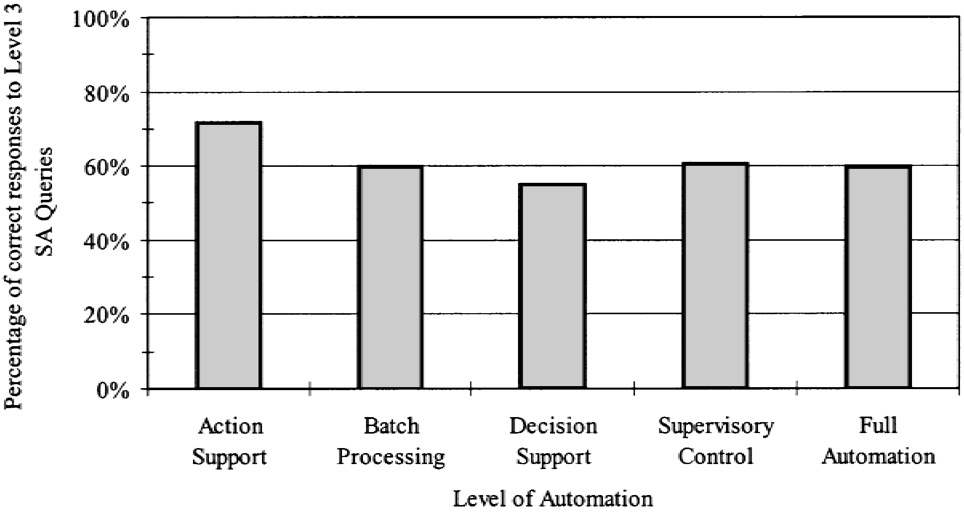


Figure 5 Mean percentage of correct responses to Level 3 SA questions as a function of LOA.

These results on operator SA are supportive of the trends in failure mode performance described earlier. SA significantly decreased with LOA and was accompanied by significant increases in time-to-recovery from failures preceded by high-level automation.

3.4. Workload

It was hypothesized at the onset of the experiment that lower levels of workload would be observed at higher LOAs that allocated the majority of task roles to the computer and removed the human operator from active processing to serve as a monitor. ANOVA results revealed LOA to be significant in its effect, $F(4,9) = 22.538$, $p < .001$, on the NASA-TLX. Figure 6 shows a plot of the mean subjective workload score as a function of LOA revealing a decreasing trend at progressively higher LOAs. Tukey's HSD tests on the overall workload scores revealed Action Support and Batch Processing to significantly differ ($p < .05$) from all other levels and from each other.

As human responsibility in the active system control loop was reduced, time-to-container completion under normal conditions decreased, demonstrating the benefits of computer processing, particularly in the implementation aspect of the task. The results on workload also reflect this reduction in involvement in the control loop. Workload significantly decreased as participants were progressively removed from plan implementation (Batch Processing), limited in their capability to plan move sequences (Decision Support), and reduced to the status of system monitor (Supervisory Control) or observer (Full Automation).

The workload results demonstrate a positive relationship with SA. A correlation analysis of Level 3 SA and the overall NASA-TLX scores across all LOAs revealed a significant positive relation ($r = 0.76$, $p < .05$). Operator Level 3 SA was significantly greater at those LOAs producing high workload. This is in line with previous research (Kaber, 1996) demonstrating greater degrees of human involvement in dynamic task control

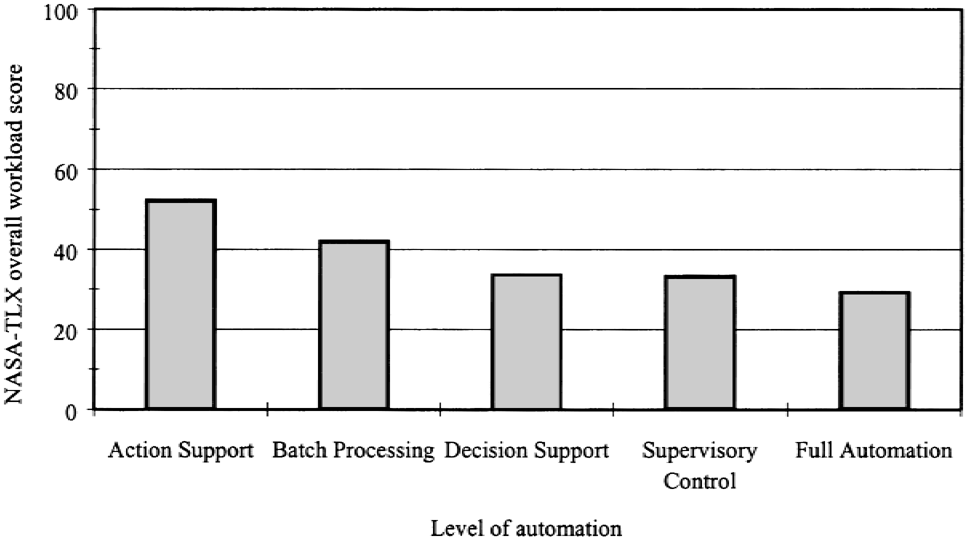


Figure 6 Mean NASA-TLX overall workload score as a function of LOA.

(through LOA) leading to an increase in subjective perceptions of workload accompanied by improved SA.

4. CONCLUSIONS

This study reviewed LOA taxonomies presented in the literature. Arguments both for and against the use of general or task-specific taxonomies were levied. In particular, a potential lack of applicability of many taxonomies to real-world tasks was raised. This issue was addressed through further empirical exploration of Endsley and Kaber's (1999) LOA taxonomy demonstrating the applicability of specific LOAs to simulated telerobot nuclear materials handling.

The experiment demonstrated that higher LOAs enhance performance during normal operating conditions through computer processing. Lower levels of subjective workload were observed at the same levels accompanying performance improvements. The experiment revealed low LOAs to promote higher operator SA and enhance human manual performance during system failure modes, as compared to high-level automation. This effect was attributed to maintaining operator involvement in the system control loop during normal operations.

Beyond the specific LOAs investigated, a combination of intermediate-level automation (Decision Support) with higher levels (e.g., Supervisory Control) produced the benefit of increased SA (in the loop control), as well as reduced workload. The off-loading of task responsibilities from operators at these levels possibly allowed for cognitive resources to be devoted to perceptual processes and thereby improved failure performance.

In general, results from this experiment confirm many of the findings of previous research (Endsley & Kaber, 1999; Endsley & Kiris, 1995) through a realistic task simulation promoting meaningfulness of the results to the design of telerobotic and general dynamic control systems. The study affirms that even when fully automated functioning may be technically possible, it may not be desirable if human-machine system performance is to be optimized across both normal operating conditions and failure modes. Results suggest that low and intermediate LOAs may not only be useful, but preferable for this purpose. However, care must be taken in generalization of these results, as a specific task type (robotic materials handling) was investigated. As well, the controlled technological limitations of the computer in the simulation had an impact on the data.

From an automation design perspective, the research validates human-centered LOA as an alternate approach to traditional automation. It provides detailed guidance to teleoperation/telerobotic systems designers who often operate under the assumption that allocating the maximum amount of system responsibility to a computer is the "best" method for safety and performance in hazardous operations.

ACKNOWLEDGMENTS

The United States Department of Energy, through a grant to the Amarillo National Resource Center for Plutonium, sponsored this work. Parts of the research were completed while the authors worked at Texas Tech University and Mississippi State University.

REFERENCES

Anderson, R.J. (1996, April). Autonomous, teleoperated, and shared control for robot systems. Proceedings of the 1996 IEEE International Conference on Robotics and Automation (pp. 2025–2032). Minneapolis: IEEE.

Billings, C.E. (1991, April). Human-centered aircraft automation: A concept and guidelines (NASA Tech. Memo. 103885). Moffet Field, CA: NASA-Ames Research Center.

Carmody, M.A., & Gluckman, J.P. (1993). Task specific effects of automation and automation failure on performance, workload and situational awareness. In R.S. Jensen & D. Neumeister (Eds.), Proceedings of the 7th International Symposium on Aviation Psychology (pp. 167–171). Columbus, OH: Department of Aviation, Ohio State University.

Draper, J.V. (1995). Teleoperators for advanced manufacturing: Applications and human factors challenges. The International Journal of Human Factors and Manufacturing, 5, 53–85.

Draper, J.V., Kaber, D.B., & Usher, J.M. (1998). Telepresence. Human Factors, 40, 354–375.

Endsley, M.R. (1987). The application of human factors to the development of expert systems for advanced cockpits. Proceedings of the Human Factors Society 31st Annual Meeting (pp. 1388–1392). Santa Monica, CA: Human Factors and Ergonomics Society.

Endsley, M.R. (1988). Design and evaluation for situation awareness enhancement, Proceedings of the Human Factors Society 32nd Annual Meeting (pp. 97–101). Santa Monica, CA: Human Factors and Ergonomics Society.

Endsley, M.R. (1995a). Measurement of situation awareness in dynamic systems. Human Factors, 37, 65–84.

Endsley, M.R. (1995b). Towards a new paradigm for automation: Designing for situation awareness. Proceedings of the 6th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design and Evaluation of Man-Machine Systems (pp. 421–426). Cambridge, MA: MIT Press.

Endsley, M.R., & Kaber, D.B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics, 42, 462–492.

Endsley, M.R., Kaber, D.B., & Onal, E. (1997). The impact of intermediate LOAs on situation awareness and performance in dynamic control systems. In D.I. Gertman, D.L. Schurman, & H.S. Blackman (Eds.), Global perspectives of human factors in power generation. Proceedings of the 1997 IEEE Sixth Conference on Human Factors and Power Plants (pp. 7/7–7/12). New York: IEEE.

Endsley, M.R., & Kiris, E.O. (1995). The out-of-the-loop performance problem and level of control in automation. Human Factors, 37, 381–394.

Hart, S.G., & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock & N. Meshkati (Eds.), Human Mental Workload (pp. 139–183). Amsterdam: North-Holland, Elsevier Science.

Kaber, D.B. (1996). The effect of level of automation and adaptive automation on performance in dynamic control environments. Doctoral dissertation, Texas Tech University, Lubbock, Texas.

Kaber, D.B., Onal, E., & Endsley, M.R. (1997). Design and development of a comprehensive user interface for teleoperator control in nuclear applications. Proceedings of the 7th Topical Meeting on Robotics and Remote Systems (pp. 947–954). LaGrange Park, IL: American Nuclear Society.

Milgram, P., Rastogi, A., & Gordski, J.J. (1995, July). Telerobotic control using augmented reality. Proceedings of the 4th IEEE International Workshop on Robot and Human Communication (RO-MAN '95). New York: IEEE.

Moray, N. (1986). Monitoring behavior and supervisory control. In K.R. Boff & J.P. Thomas (Eds.), Handbook of perception and human performance: Vol. II. Cognitive processes and performance (pp. 40–43). New York: John Wiley & Sons, Inc.

Parasuraman, R., Molloy, R., & Singh, I.L. (1993). Performance consequences of automation induced complacency. International Journal of Aviation Psychology, 3, 1–23.

Sheridan, T.B. (1992). Telerobotics, automation, and human supervisory control (2nd ed.). Cambridge, MA: MIT Press.

Wickens, C.D. (1992). Engineering psychology and human performance (2nd ed.). New York: HarperCollins.

Wiener, E.L. (1988). Cockpit automation. In E.L. Wiener & D.C. Nagel (Eds.), Human factors in aviation (pp. 433–459). San Diego, CA: Academic Press.