Effects of Automation of Information-Processing Functions on Teamwork

Melanie C. Wright, Duke University Medical Center, Durham, North Carolina, and **David B. Kaber,** University of Rostock, Rostock-Warnemünde, Germany

We investigated the effects of automation as applied to different stages of information processing on team performance in a complex decision-making task. Forty teams of 2 individuals performed a simulated Theater Defense Task. Four automation conditions were simulated with computer assistance applied to realistic combinations of information acquisition, information analysis, and decision selection functions across two levels of task difficulty. Multiple measures of team effectiveness and team coordination were used. Results indicated different forms of automation have different effects on teamwork. Compared with a baseline condition, an increase in automation of information acquisition led to an increase in the ratio of information transferred to information requested; an increase in automation of information analysis resulted in higher team coordination ratings; and automation of decision selection led to better team effectiveness under low levels of task difficulty but at the cost of higher workload. The results support the use of early and intermediate forms of automation related to acquisition and analysis of information in the design of team tasks. Decision-making automation may provide benefits in more limited contexts. Applications of this research include the design and evaluation of automation in team environments.

INTRODUCTION

A large number of work environments that use automation are so complicated that they require multiple operators to simultaneously address tasks and manage automation (Bowers, Oser, Salas, & Cannon-Bowers, 1996). Although researchers have noted human-automation interaction problems associated with the design of automation in aviation and other domains (Bainbridge, 1987; Coury & Semmel, 1996; Woods, 1996), most of this research has focused on the effects of automation on an individual worker. Some research has suggested that automation may qualitatively change communication between human team members (Johannesen, Cook, & Woods, 1994; Wiener, 1993). With this in mind, there is a need to ascertain the effects of automation on teamwork in order to promote the design of safe and effective systems.

Prior research assessing the effect of automation on teams has produced contradictory results. We first review this research and offer possible explanations for the variety of results. We then suggest an alternative approach for assessing the effect of automation on teams by classifying automation based on its application to different human-machine system information-processing functions in accordance with existing theories of levels of automation (LOAs). In order to predict the effect of different LOAs on teams, we also review research assessing the effects of LOAs on the performance of individuals and research describing characteristics of high-performing teams. The purpose of this study was to determine whether differences in the form of complex automation have implications for team coordination and performance, to explain these effects in terms of the functional nature of the automation, and to establish how automation may mediate the potential for coordinated teams to achieve high performance.

Address correspondence to Melanie Wright, Department of Anesthesiology, Duke University Medical Center, Box 3094, Durham, NC, 27710; melanie.wright@duke.edu. **HUMAN FACTORS,** Vol. 47, No. 1, Spring 2005, pp. 50–66.

Effects of Automation on Teams

Existing research on flight deck automation and teams (Bowers, Deaton, Oser, Prince, & Kolb, 1993; Clothier, 1991; Costley, Johnson, & Lawson, 1989; Wise, Guide, Abbot, & Ryan, 1992) provides support for the notion that automation has some type of effect on team coordination; however, results are mixed. Some studies show increases in communication rates (Wise et al.), some show rate decreases (Costley et al.), some suggest detriments to team coordination (Bowers, Deaton, et al.), and some suggest coordination improvements (Clothier). One consistent result across studies on verbal communication and automation is that significant team coordination differences between systems tend to appear when workload is higher (Clothier; Costley).

A possible reason for the conflicting findings is that studies generally compare team coordination under a new automated system with teamwork in an earlier model aircraft (either through surveys, field studies, or comparisons in high-fidelity simulators); therefore other differences, such as advanced displays, may be affecting team coordination. A second potential problem is that different forms of automation may influence team coordination in different ways. Bowers, Jentsch, and Salas (1994) offered some solutions to these problems, including (a) studying automation using low-fidelity simulation in a laboratory environment so that differences between conditions can be isolated to only those associated with automation and (b) clearly specifying the form of automation used – for example, through a taxonomy of automation.

Jentsch and Bowers (1996) applied these recommendations using a PC-based flight simulator in which the pilot and copilot performed with a single form of automation (autopilot and navigation computer, respectively) on or off. The autopilot represented automation of psychomotor tasks, whereas the navigation computer represented automation of more cognitive functions. They found that performance improved when both automated systems were in use. They also found decreased rates of communication associated with the use of the navigation computer (copilot automation), possibly because the copilot spent time interacting with the navigation computer rather than coordinating with the pilot.

Levels of Automation

Although Bowers et al. (1994) offered a taxonomy of automation for aviation, more general taxonomies and models have been developed to describe different forms of automation. Scerbo (1996) provided a review of taxonomies and models of LOAs that have evolved from the early work of Sheridan and Verplanck (1978). Sheridan and Verplanck's taxonomy distinguished 10 different LOAs that focused on who has decision authority in system operations (the human or a computer), what information is provided to the user by the system, and who implements an action.

More recent taxonomies and models consider automation as applied to a wider range of functions (Endsley, 1997; Endsley & Kaber, 1999; Kaber & Endsley, 1997; Parasuraman, Sheridan, & Wickens, 2000; Wickens, Mavor, Parasuraman, & McGee, 1998). The taxonomy developed by Endsley and Kaber demonstrates how automation can be classified in terms of four cognitive and psychomotor aspects of human information processing: monitoring, generating, selecting, and implementing. Based on an analysis of real-world systems, they established 10 specific LOAs depending on whether or not the human or a computer performed each of the four information-processing functions or if the functions were shared. The function allocation scheme also defined the overall level of system autonomy.

Parasuraman et al. (2000) developed their model based on the premise that automation is a continuum and not an "all-or-none concept." Like Endsley and Kaber (1999), they identified four stages of human information processing that can be used as a basis for classifying automation: information acquisition, information analysis, decision selection, and action implementation. They said that automation can vary from low to high along unique continua for each information-processing function. The Parasuraman et al. model is a theoretical one that can be used to describe virtually any human-machine system in terms of the type and level of automation. Approaches to classifying LOAs, such as these, serve to qualitatively define the characteristics of automation systems so that they may be compared and evaluated in research.

Effects of LOAs on Individuals

Researchers have found that, similar to the goals of adaptive automation, providing the right degree of automation for the right function of a task can optimize the use of automation in terms of operator workload and situation awareness (SA; Endsley & Kaber, 1999; Endsley & Kiris, 1995; Kaber & Endsley, 1997; Parasuraman et al., 2000). Comparisons have been made of the effects of various LOAs on individual operator performance. Endsley and Kaber found that levels of automation that combine human generation of options with computer implementation of actions produced better overall performance during normal operations of a laboratory simulation of an air traffic control task. In a visual target identification task, Galster, Bolia, and Parasuraman (2002) found performance advantages associated with information automation (cuing) that were not apparent when automation was applied to decision making.

Kaber, Onal, and Endsley (2000), using a high-fidelity simulation of a telerobot, found that high levels of automation involving computer assistance in information analysis and action implementation, or assistance in these functions plus decision making, enhanced performance and reduced workload during normal operation conditions. Intermediate levels of automation including computer assistance in action implementation promoted higher operator SA and enhanced manual performance during system failure modes, as compared with higher levels of automation. Clamann, Wright, and Kaber (2002) found that operators were better able to adapt to adaptive automation when the automation was applied to information acquisition and action implementation than when automation was applied to information analysis and decision-making tasks. Rovira, Zinni, and Parasuraman (2002) found that automation unreliability had a greater cost for decision automation than for information automation. Sarter and Schroeder (2001) obtained the same results in investigating information and decision automation in pilot assessment of in-flight icing conditions.

In general, this body of work supports automation of psychomotor functions (e.g., automation of information acquisition and action implementation) and information automation as compared with automation of higher-order cognitive functions such as decision selection, particularly in situations in which the automation may be unreliable, adaptive automation is being used, or operators are exposed to high workload.

Characteristics of High-Performing Teams

Studies of team communication and coordination have noted specific types of behavior that are associated with good team performance (Costley et al., 1989; Foushee, Lauber, Baetge, & Acomb, 1986; Orasanu, 1990). It has been observed that teams in which members provide unsolicited information to other team members generally perform better than those that do not (Johannesen et al., 1994; Urban, Bowers, Monday, & Morgan, 1993). High-performing teams tend to be more efficient in their use of questions, asking fewer questions yet still receiving all the necessary information (Urban et al., 1993). High-performing teams also exhibit behaviors such as situation assessment and planning that help to achieve and maintain SA (Orasanu).

Jentsch, Sellin-Wolters, Bowers, and Salas (1995) found that teams that were faster in detecting a problem used more standard communications, made more leadership statements, and vocalized more SA observations than did slow teams. Bolstad and Endsley (1999, 2000) found that a training condition that allowed for the development of similar mental models through the use of shared information displays led to improved performance in a decision-making task. MacMillan, Entin, and Serfaty (2002) suggested that although there is a cost associated with a "need for communication" between team members, there is a benefit when communication is in the form of collaborative planning. It is not yet clear whether the implementation of certain forms of automation, as defined in the literature, may promote or support behaviors that are frequently associated with good team performance.

OBJECTIVE AND HYPOTHESES

The present study evaluated the effect of automation on teamwork in a decision-making task. In order to more completely describe the effects of automation, we designed the task to present three different levels of automation: computerization of (a) information acquisition, (b) information analysis, and (c) decision selection functions as defined in the model presented by Parasuraman et al. (2000). Four experimental conditions, designed to represent realistic forms of automation, were mapped to conditions presented in Endsley and Kaber's (1999) taxonomy of automation: (a) action support, (b) shared control, (c) decision support, and (d) blended decision making.

The action support condition represents the lowest level of automation across the three functions manipulated and provides a baseline automation condition. Action support is representative of a system that assists the operator with performance of an action, with some human control required. An example is a teleoperation system, in which the behavior of the robot is directly slaved to human inputs. The other conditions represent increases in information acquisition (shared control), information analysis (decision support), and both information analysis and decision selection (blended decision making) from the baseline condition (see Table 1). With shared control, which is a more sophisticated form of automation, the human retains decision authority regarding the system behavior and the automation implements the operator's plan. The operator can also modify the system behavior in real time. Decision support is representative of a system integrating an expert or decision support system. The operator may exploit or ignore this capability. Blended decision making is representative of a system integrating a high-level decision support aid capable of automatically selecting among decision alternatives and implementing specific behaviors. Under this mode of automation, the human can consent to automated decisions and actions or override them.

The objective of this study was to assess the effects of different types and levels of automation on the performance and coordination of teams and, based on this analysis, to identify ways in which human-automation interaction can be improved in a team environment. The results of team performance studies suggest that there are three main ways in which automation conditions can facilitate teamwork: (a) by reducing workload (particularly as it relates to the "need for communication") – lower workload will provide teams with additional time for planning, situation assessment, and coordination efforts; (b) by promoting operator SA – individuals with higher levels of SA will be better able to predict the needs of their team members, resulting in more efficient communications, and may make more efforts at vocalizing SA; and (c) by supporting the development of shared mental models (e.g., through the use of shared displays) – teams with shared mental models are also more likely to exhibit more efficient communications and will be facilitated in their team planning and coordination efforts.

Research on the effects of automation on individual performance (Clamann et al., 2002; Galster et al., 2002) suggests that automation of information acquisition functions may facilitate teamwork through all three of these mechanisms. Consequently, we hypothesized that this type of automation (shared control) would lead to more efficient team communications (H1), a greater incidence of team coordination behaviors such as planning and situation assessment (H2), and better team effectiveness with respect to the task (H3), when compared with a baseline condition (action support). Intermediate levels of automation, such as automation of information analysis functions (decision support), may also have positive effects on teamwork through workload reductions and SA improvements (Kaber et al., 2000). Thus we expected communication quality (H4), team coordination

TABLE 1: Level of Automation Comparisons

Comparison	Increasing Automation of:
Action support vs. shared control	Information acquisition
Action support vs. decision support	Information analysis
Action support vs. blended decision making	Information analysis and decision selection
Decision support vs. blended decision making	Decision selection

(H5), and team effectiveness (H6) to be improved by this condition as compared with the baseline condition.

Decision automation, however, has not been associated with increases in SA or workload reductions (Clamann et al., 2002; Galster et al., 2002; Rovira et al., 2002). Therefore, in a comparison between a condition with both information analysis and decision automation (blended decision making) and one providing only information analysis automation (decision support), we expected communication quality (H7), team coordination (H8), and team effectiveness (H9) to be better when decision automation was not present. Finally, we hypothesized that these differences between conditions would be more pronounced under conditions of high task difficulty than under low task difficulty (H10).

METHOD

Task

The team performance task was based on a laboratory simulation of the Theater Defense Task (TDT), which was developed by Bolstad and Endsley (1999) based on an individual-control task developed by Kaber and Endsley (1997). The TDT is a team decision-making and target elimination task completed by two team members. The intelligence officer (IO) and air commander (AC) have separate but interrelated tasks. They work at separate workstations connected by a local area network. The role of the AC is to protect a home base from incoming enemy aircraft. The role of the IO is to classify incoming targets as enemy or friendly and to indicate the type of aircraft for the AC. Based on the aircraft type, the AC must either choose an appropriate missile to destroy enemy aircraft or allow friendly aircraft to pass through.

The task is relatively complex in that it requires the IO to consider several data sources simultaneously, and to account for the sensitivity of the sources, in determining the classification of a target. The task requires communication between the AC and the IO, some of which occurs via the computer interfaces, such as the IO sending classification information to the AC's display. In addition, verbal communication is required from the AC to the IO regarding target positions and airborne warning and control system (AWACS) reconnaissance aircraft positions (or sensor reliability). The AWACS aircraft travel around the airspace in a random pattern, providing greater reliability when they are closer to home base. The pace of the task is such that both team members must be strategic about how they prioritize and process targets in order to achieve a high score. Figure 1 presents the AC's task display, and Figure 2 presents the IO's task display.

A team receives points when targets are destroyed or when targets reach the home base. Damage points are assessed when an enemy aircraft reaches the home base. Target points are given when an aircraft is destroyed. Positive target points (rewards) are earned for destroying enemies, whereas negative target points (penalties) are assessed for destroying friendly aircrafts. The number of points assigned is dependent on the mission relevance and the lethality of the aircraft.

Levels of automation. Manipulations of automation along the first three of the four functions included in the Parasuraman et al. (2000) model (information acquisition, information analysis, and decision selection) were selected as they were expected to yield the greatest differences in team coordination based on prior research (whereas changes in automation of action implementation were not expected to greatly influence teamwork). This selection provides experimental comparisons for studying increases in each of the three types of automation and an increase in both information analysis and decision support (see Table 1). Modifications to the AC's and IO's displays used under the baseline condition (action support) in order to model the three other automation conditions are presented in Table 2.

Experimental Design

A $4 \times (2 \times 2)$ mixed design was used in this experiment, with the four LOAs representing a between-subjects variable and the two levels of task difficulty and two order conditions representing within-subjects variables. Participants were randomly assigned to automation conditions and team member roles. Level of difficulty was determined by the number of targets present on the task displays at any given time (five targets for a low-difficulty condition and eight

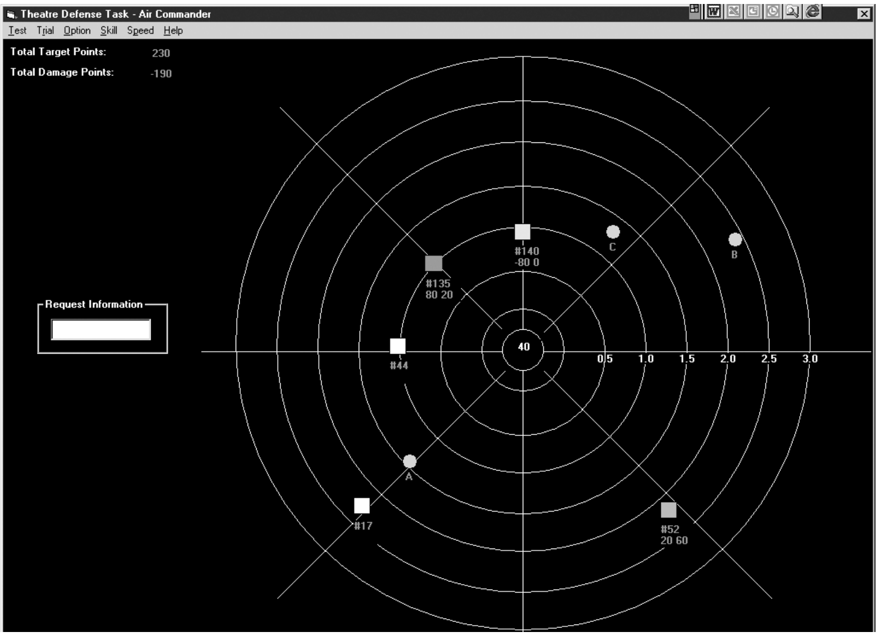

Figure 1. Air commander's display.

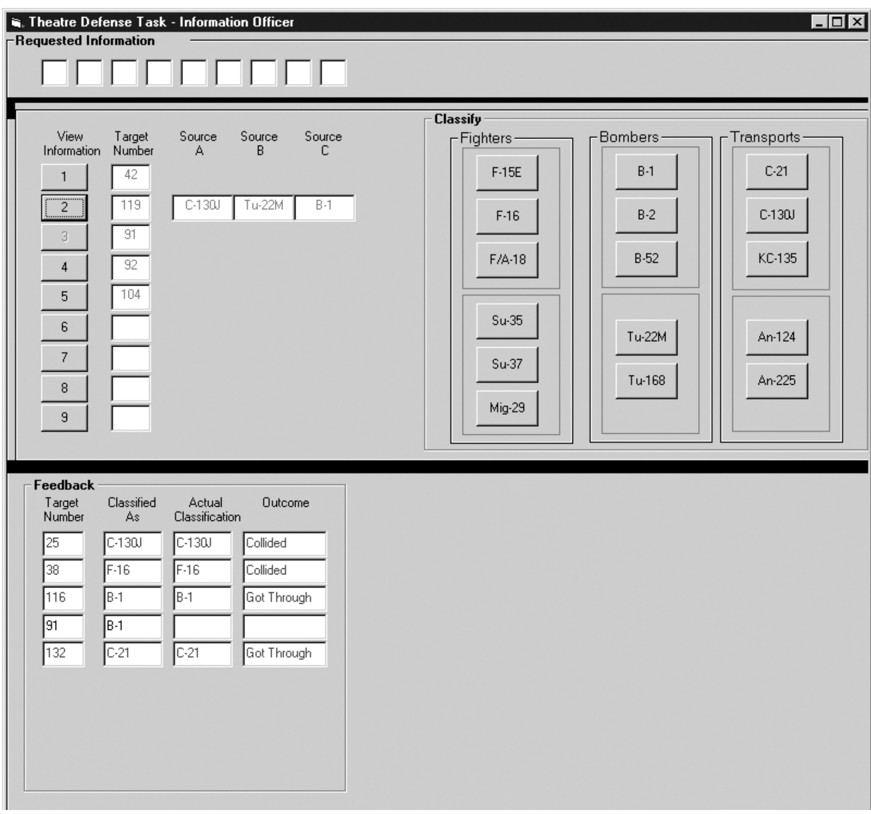

Figure 2. Intelligence officer's display.

TABLE 2: Changes to Theater Defense Task to Implement Automation Conditions

Condition	Changes to Baseline (Action Support) Condition
Shared control	For automation of information acquisition function, task-relevant information for IO was abstracted from AC's display and presented to IO and vice versa (based on shared displays of Bolstad & Endsley, 1999, 2000). For IO, information included target proximity, sensor reliabilities, and points scored. For AC, information included sensor classification of targets, after IO selected a classification. Target proximity presentation on IO display was also considered a minor increase in information analysis automation as it provided a possible priority order for classifying targets.
Decision support	Information analysis automation was implemented by presenting IO target list in order indicating target that should be classified first based on distance to home base and target type information provided by sensors (e.g., enemy targets close to home base presented first in list). List was continually re-sorted each time target reached center of display or was eliminated.
Blended decision making	In addition to priority sorted target list provided in decision support condition, decision selection automation involved modification of IO display to include fourth column in target/sensor data section (to right of Source C column in Figure 2) displaying automated target classification decision. Algorithm for target classification was at least as good as human team, except when all sensor reports had equal reliability. In this case, human team could consider additional information (e.g., risk of reward or penalty associated with decision) that algorithm did not take into account.

targets for a high-difficulty condition). Teams were presented with two trials under each difficulty condition. Two orders of task difficulty (across trials) were used in the study: low-high-high-low and high-low-low-high. These settings of order were balanced across automation conditions and considered as an additional nested variable in the experimental design.

Participants and Apparatus

Ten teams of 2 persons were assigned to each automation condition for a total of 40 teams (80 participants). This sample size was selected based on a power analysis on a pilot data set with significant effect size estimates. The participants were students at North Carolina State University who participated for monetary compensation. All were required to have 20/20 or corrected-to-normal vision and some PC experience with a graphical user interface. There were 17 male-male teams, 14 male-female teams, and 9 female-female teams. Teams were randomly assigned to automation condition with gender composition partially balanced across the four automation conditions (either 4 or 5 male-male teams; 3 or 4 male-female teams; and 2 or 3 female-female teams in each automation condition). The average ages of the ACs and IOs were 26.5 and 24.6 years, respectively.

Team members sat at adjacent workstations separated by a partition so that they could not see each other. Team members communicated with each other using FM radio headsets (walkie-talkies) or through the computer interfaces. The headset covered one ear, and a "press-to-talk" button was used for transmitting. The same communication methods were available under all forms of automation. A digital video camera recorded the ongoing task activity on the two computer displays and test participant communications.

Dependent Measures

Team effectiveness measures. Team effectiveness was measured using the TDT point-scoring system. In addition to target points and damage points, a total score was determined by subtracting the damage points from the target points. Because there is a greater scoring opportunity when more targets are presented (in the high-difficulty condition), team effectiveness was also measured as a ratio of target points, damage points, and total score to the number of targets presented. The target points ratio provided a measure of the team's decision-making effectiveness for a single target. The damage points ratio provided a measure of the team's ability to keep pace with the task and, to a lesser extent, the team's decision making.

Team communication counts measure. One method of measuring team communication is to categorize communication types and then quantify communication in accordance with the categorization (Costley et al., 1989; Urban et al., 1993). This type of measure allowed us to quantify specific task-relevant communication and relate these counts to differences in performance or team effectiveness. Pilot testing and analysis of transcripts were used to identify the types of task-relevant communications that team members used in the TDT. These were reduced to the following seven communication types (for analysis): (a) target classification request from AC with response from IO; (b) target classification request from AC without response from IO; (c) unsolicited target classification reports by IO; (d) sensor reliability requests from IO with response from AC; (e) unsolicited sensor reliability update from AC; (f) other communications; and (g) total communications.

Target classification requests refer to the calling out of specific target numbers by the AC for classification by the IO. A classification request without a response does not necessarily mean the IO ignored the request from the AC, only that he or she made no verbal response. In most cases this indicated that the IO chose to respond to the AC using the user interface (by color coding the targets) rather than responding verbally. The category of other communications includes pretrial strategy comments, strategy comments during the trial, and other communications, such as sharing information regarding the current score or decision-making discussion. Two raters counted the communications during a trial by marking a tally for each type of utterance.

The communication counts were also used to calculate an anticipation ratio (Serfaty, Entin, & Johnston, 1998) by summing the tallies that represented information transfer between team members, summing the tallies that represented information requests, and then calculating the ratio of the information transferred to the information requested. This measure provides an indication of the degree to which individuals predicted the information needs of their team members and can be related to the amount of solicited versus unsolicited information transfer.

Team coordination measure. The measurement of team coordination used in this study followed the conceptual framework of Dickinson and McIntyre (1997). This framework proposes the development of observation (or event) scales that indicate specific behaviors associated with high or low performance along several dimensions of teamwork. The scales are then used to record team performance on these behaviors as a basis for a final team rating. The rating is made on a 5-point scale for each teamwork dimension. The teamwork dimensions used in this study were similar to those used by Brannick, Prince, Prince, and Salas (1995) and a subset of those defined for the aircrew coordination observation and evaluation scale (Bowers, Morgan, Salas, & Prince, 1993). They were (a) assertiveness, (b) decision making, (c) situation assessment, (d) leadership, and (e) communication.

In using this method, two raters recorded observations of good or poor teamwork on each teamwork dimension. At the end of a trial, raters considered both the number of times specific teamwork behaviors were exhibited, as well as the quality of the behaviors, in their final ratings on each of the five dimensions of teamwork. Raters used a scale from 1 (*hardly any skill*) to 5 (*complete skill*) for each dimension rating, and a total team coordination rating was determined by summing the 5-point ratings on each of the dimensions of teamwork.

Subjective workload measure. Both team members rated perceived workload. They were asked to complete rankings and ratings of various workload dimensions using the modified NASA-Task Load Index (TLX; Hart & Staveland, 1988).

Procedure

The procedures for the experiment consisted of five steps: (a) introduction, completion of consent forms and background questionnaire; (b) training on the TDT; (c) a 15-min practice session in the low-difficulty condition; (d) completion of a practice rating of workload using the NASA-TLX; and (e) four 15-min trials, each followed by a workload rating.

Ten-minute breaks were given following the practice trial and between test trials. The duration of the practice session was based on the time required to achieve asymptotic performance in the task as determined by prior research (Kaber & Riley, 1999). The entire experiment lasted between 2.5 and 3 hr.

An extensive training session was given to the team members verbally by the experimenter. The barrier between them was removed, and both team members received instructions on the entire task, including the role of both the AC and the IO. Team members were allowed 5 min to try out the task in their role, using the radios to communicate. For the formal TDT practice session and the experimental trials, the barrier was placed between team members and they were asked to imagine that they were in different locations and to only use the radios for communication. They were allowed 5 min to discuss (via radio) their team strategy for performing the task prior to the start of the practice session and the trials.

DATA ANALYSIS

Team Effectiveness and Workload

For the evaluation of the team effectiveness and workload measures, a three-way analysis of variance (ANOVA) model was used with difficulty as a within-subjects variable and automation condition and order as between-subjects variables. In the cases in which this model did not result in any significant effects attributable to order, analyses were conducted with a reduced model including only task difficulty and LOA as variables. Significant effects revealed by the ANOVAs were further investigated using Duncan's multiple range (MR) test with an alpha level of .05.

Because of the nature of the TDT, sensor reliability was randomly generated during the trial and could not be controlled. In order to assess the potential impact of sensor reliability on performance measures, we conducted correlation analyses on the average sensor reliability over a trial and each of the team effectiveness measures. All team effectiveness measures correlated significantly ($p < .05$) with sensor reliability. Therefore the team effectiveness measures were analyzed using the ANOVA models described previously with the addition of a covariate, overall sensor reliability, as a random effect in order to control for the influence of this factor.

Team Communication Counts and Anticipation Ratio

Interrater reliabilities (correlations) were calculated for the communication counts. Significant reliabilities ($p < .0001$) were obtained for counts on all seven communication types, and all correlation coefficients were greater than $r = .9$ with the exception of other communications, which had a coefficient of .79.

To determine if any of the specific types of communications tallied were associated with either high or low performance, we conducted linear correlations on the three effectiveness measures, each of the seven communication types, and the anticipation ratio. There were 160 observations (40 teams × 4 trials) on each measure for this analysis.

For the analysis of communication counts by automation condition, the communication counts failed to meet the normality assumptions of the ANOVA. For some of the counts there was a strong floor effect (many observations of zero), and transformations of the data did not resolve the ANOVA assumption violations. Therefore, both the counts of individual types of communication data and the anticipation ratio (calculated from communication count data) were subject to nonparametric analyses, based on ranks.

For effects attributable to automation condition, the Kruskal-Wallis test was used and the data were divided into two separate sets for analysis, one representing the low-difficulty condition and the other representing the high-difficulty condition. In addition, the data were averaged across the two trials under each difficulty condition and across the two raters. For those measures for which the result was significant, separate Kruskal-Wallis tests were conducted on pairs of automation conditions (e.g., action support vs. shared control) to determine which of the conditions were significantly different from each other. For effects attributable to level of difficulty, a two-tailed Wilcoxon signed ranks test was used. The data were separated into four sets, one for each automation condition, and were averaged across the two trials and the two raters.

Team Coordination Measure

Interrater reliabilities (correlations) were calculated for the team coordination rating. A significant ($p < .0001$) interrater correlation was obtained ($r = .56$). Correlations between raters in other studies using this method have ranged from $r = .55$ to $.87$ (Brannick, Roach, & Salas, 1993), $r = .48$ to $.71$ (Travillian, Volpe, Cannon-Bowers, & Salas, 1993), and $r = .55$ to $.97$ (Volpe, Cannon-Bowers, Salas, & Spector, 1996). The team coordination ratings failed to meet the normality assumptions of the ANOVA because of the discrete nature of the data (ratings of 1–5 on each dimension). The application of transforms to the response was not successful in accounting for the assumption violation, so the team coordination ratings were subjected to the same nonparametric tests as described for the communication counts.

RESULTS

Difficulty Manipulation Check

There were significant effects for four of the six team effectiveness measures, with a greater number of target points, $F(1, 36) = 27$, $p < .0001$, and damage points, $F(1, 32) = 81$, $p < .0001$, in the high-difficulty condition compared with the low-difficulty condition because of the higher number of targets in the former condition. In addition, there was a higher total score ratio, $F(1, 36) = 98$, $p < .05$, and a lower damage ratio, $F(1, 32) = 21$, $p < .0001$, in the low-difficulty condition compared with the high-difficulty condition. Participants were better able to keep pace with the task and were better able to classify and resolve targets correctly on a per-target basis in the low-difficulty condition than in the high-difficulty condition. Ratings of workload were higher in the high-difficulty condition than in the low-difficulty condition for both the AC role, $F(1, 32) = 19.1$, $p < .0001$, and the IO role, $F(1, 36) = 63.9$, $p < .0001$. (The denominator degrees of freedom for the F tests varied across response measures as a result of the order effect being retained in, or removed from, the statistical model.)

The only significant effect of task difficulty on any of the team coordination measures was in the counts of "other communications." The Wilcoxon signed ranks test indicated that for the blended decision-making condition, there were significantly more "other communications" under the low-difficulty condition than in the high-difficulty condition ($Z = 2.28$, $p < .05$; nonparametric tests were used to analyze team communication counts and coordination ratings). This indicates that the low-difficulty condition may have allowed more time for communications other than task-relevant communications in this automation condition.

Correlations Between Communication Counts and Effectiveness Measures

The significant correlation analyses on the communication counts and team effectiveness measures are shown in Table 3. Because lower damage points and damage ratios indicate better performance, a negative correlation of these measures with any communication type indicates that the communication was associated

TABLE 3: Correlations of Communication Types and Team Effectiveness Measures

	Pearson Correlation Coefficient, r, $N = 160$ Prob. > |r| Under H_0: $\rho = 0$, p			
Communication Type	Target Points	Damage Ratio	Total Points	Total Score Ratio
Classification request from AC with response from IO	–.27** <.0005	.27** <.0005	–.29** <.0005	–.23** <.005
Unsolicited classification from the IO	.006 <.94	–.17* <.04	–.08 <.33	.01 <.90
Other communications	.14 <.08	–.11 <.15	.16* <.05	.14 <.07

*$p < .05$. **$p < .01$.

with higher performance. Results revealed that classification requests with response tended to be associated with lower performance (negative correlation with target points, total points, and total score ratio, and positive correlation with damage ratio). Unsolicited classifications were associated with better performance measured in terms of the damage ratio. Other communications were associated with better performance in terms of total points. These findings are consistent with team performance research showing that teams that use more unsolicited communications, and teams that communicate more in relation to planning and situation assessment updates (included in the counts of "other" communications here), tend to perform better than teams that do not (Johannesen et al., 1994; Macmillan et al., 2002; Urban et al., 1993).

Automation Condition Effects

Although the data analysis included comparisons of all combinations of the automation conditions, only the significant effects relevant to the hypotheses of the study (comparisons of shared control and decision support against action support, and comparisons between decision support and blended decision making) are reported here.

Communication counts. The Kruskal-Wallis tests indicated significant effects of automation condition for unsolicited reliability reports under both low difficulty, $T(3) = 15.8$, $p < .01$, and high difficulty, $T(3) = 14.4$, $p < .01$. Additional Kruskal-Wallis tests conducted on pairs of the automation conditions revealed that there were fewer unsolicited reliability reports when shared control was used as compared with action support under both the low task difficulty condition, $T(1) = 8.7$, $p < .01$, and the high task difficulty condition, $T(1) = 8.3$, $p < .01$. This was probably attributable to the reduced need for communication of reliability information from the AC to the IO under the shared control condition because the reliability information was included on the IO's display. This provides support for the assertion that information automation can reduce the need for communication, which may result in a workload reduction. There were no significant differences between automation conditions related to the three types of communications that correlated with team effectiveness measures. However, there was a trend to suggest that under high difficulty, teams in the decision support condition verbalized more "other" communications than did teams in the blended decision-making condition, $T(3) = 6.7$, $p < .10$. Although not statistically significant, this provides some support for our assertion that team coordination would be better when decision automation was not included (H7, H8).

Anticipation ratio. Figure 3 presents the mean anticipation ratio for the four automation conditions. A higher anticipation ratio indicates that more information was transferred as compared with the information that was requested. The Kruskal-Wallis test revealed a significant effect of automation under both the low-difficulty, $T(3) = 9.2$, $p < .05$, and high-difficulty conditions, $T(3) = 9.2$, $p < .05$. The comparison of the automation conditions revealed that shared control had a higher anticipation ratio than did action support for both the low-difficulty condition, $T(1) = 6.2$, $p < .05$, and the high-difficulty condition, $T(1) = 5.1$, $p < .05$. This supports our hypothesis (H1) that information acquisition automation will lead to more efficient communications than will a baseline action support condition. Automation of information analysis did not reveal more efficient communications as determined by the anticipation ratio measure, as was expected (H4 and H7).

Team coordination ratings. Figure 4 presents the average total rating of team coordination for each of the four automation conditions under the two levels of difficulty (higher ratings indicate higher teamwork skills). The Kruskal-Wallis test on total team coordination indicated significant effects of automation under both low task difficulty, $T(3) = 8.9$, $p < .05$, and high difficulty, $T(3) = 9.8$, $p < .05$. Kruskal-Wallis tests conducted on pairs of automation conditions indicated that under the low task difficulty setting, the decision support condition received higher ratings than did action support, $T(1) = 5.0$, $p < .05$, and blended decision making, $T(1) = 7.2$, $p < .01$. Under the high task difficulty setting, the decision support condition received higher ratings than did blended decision making, $T(1) = 8.1$, $p < .01$. These results support our hypothesis that information analysis automation would lead to improved team coordination over both the baseline condition (H5) and a condition including decision automation (H8).

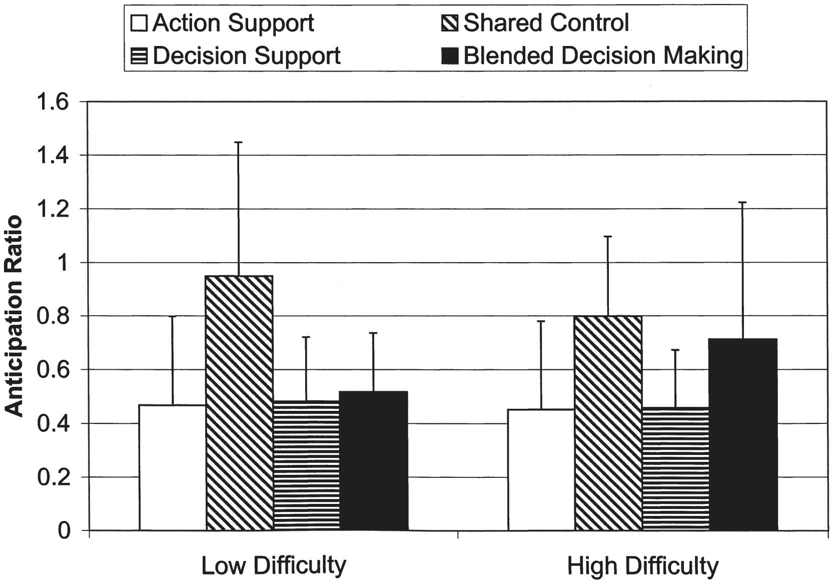

Figure 3. Anticipation ratio by automation condition.

The hypothesis that information acquisition automation would lead to improved team coordination (H2) was not supported by the team coordination measure.

Automation and Task Difficulty Interactions

Team effectiveness. There was a significant interaction between the automation condition and level of difficulty in terms of the damage points ratio, $F(3, 32) = 3.94$, $p < .05$ (see Figure 5; lower ratio values indicate higher effectiveness). The results of Duncan's MR test on the interaction indicated that the blended decision-making condition led to significantly better performance under low difficulty as compared with the action support condition with respect to allowing enemy aircraft to reach the home base. This performance enhancement of the blended decision-making condition was not present under high difficulty. This result was counter to our hypotheses regarding the effects of levels of automation on team effectiveness (H3, H6, H9). However, trends across both low- and high-difficulty conditions for each of the team effectiveness measures suggested an advantage for the information analysis automation as part of the decision support condition (in line with hypotheses H6 and H9).

Workload measures. There was a significant interaction of automation condition and level of difficulty on the IO workload ratings, $F(3, 36) = 3.5$, $p < .05$ (see Figure 6). The results of Duncan's MR test on the interaction revealed that under both low and high levels of difficulty, the blended decision-making condition received higher workload ratings than did action support and decision support. In addition, under low difficulty only, the decision support condition received higher workload ratings than did the action support condition. These results are consistent with other research findings that some forms of automation, particularly decision automation, may lead to increases in workload (Clamann et al., 2002; Galster et al., 2002; Rovira et al., 2002), which we expected to negatively affect team coordination.

DISCUSSION

The goal of this experiment was to assess the effect of different forms of automation on the performance and coordination of teams. The results provide some explanation for the conflicting findings with respect to automation and teamwork. Different forms of automation in this experiment resulted in different effects on teamwork. The automation of information acquisition

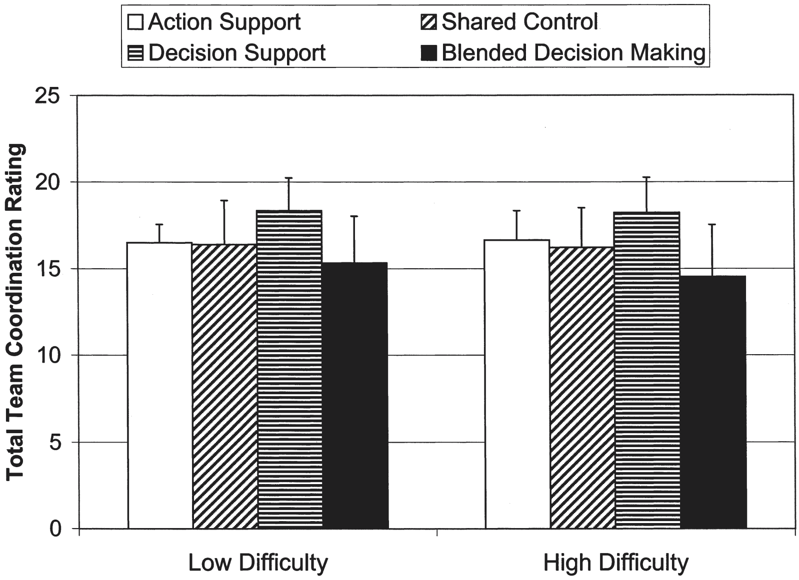

Figure 4. Average rating of total team coordination across automation conditions.

changed the communication patterns of teams, as evidenced through differences in counts of task-specific communications and in the anticipation ratio of communications, although this did not affect ratings of their team coordination skills or their effectiveness in the task.

The automation of information analysis led to higher ratings of teamwork skills. This is probably a result of a change in task strategy for some of the teams. The automated target sort encouraged teams to change from a strategy in which the AC identified the order of addressing targets to a more efficient strategy in which the IO selected the order (reducing visual search time on the part of the IO). The significant improvement in team coordination ratings and a trend

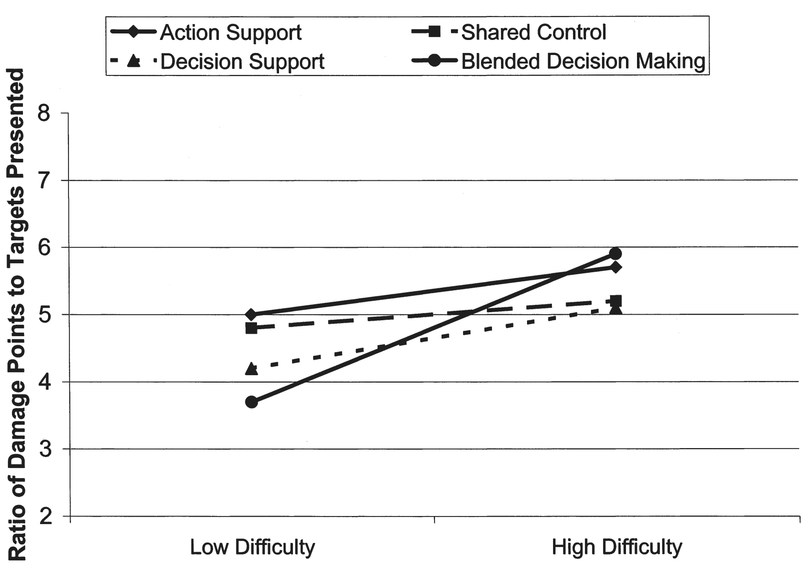

Figure 5. Damage point ratio by automation condition and level of difficulty.

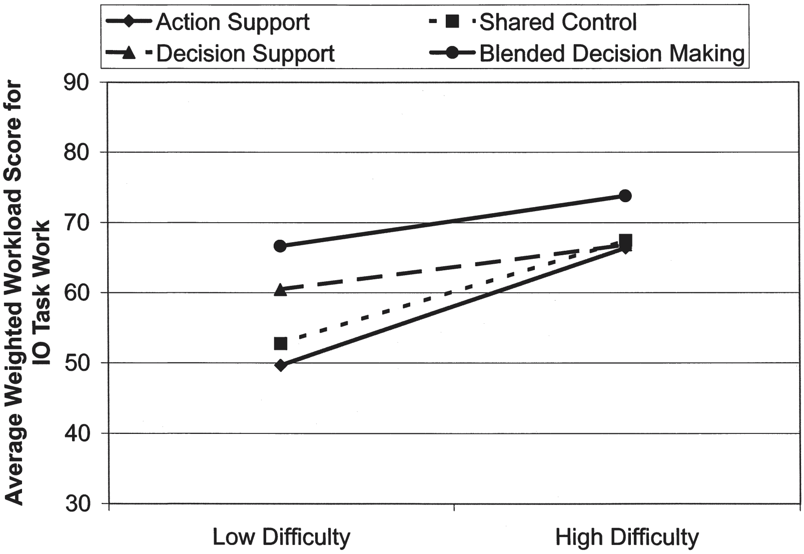

Figure 6. Automation condition by level of difficulty interaction for IO rating of task work.

toward higher counts of "other" communications suggest that this strategy change allowed the team to devote more time to communications that are indicative of greater teamwork skill (such as the "collaborative planning" communications described by MacMillan et al., 2002). Although these improvements in team coordination were not reflected in significant differences in team effectiveness measures, the decision support condition did have a trend toward the highest team effectiveness scores overall.

Decision automation negatively affected team coordination ratings, as expected. The higher workload associated with this condition left little time for teams to exhibit behaviors that represented good decision making, leadership, or situation assessment. Counter to our hypotheses, the additional information provided by this automation condition allowed team members to perform more effectively under low task difficulty conditions, at a cost of higher workload. This finding suggests that advanced forms of automation providing information analysis and decision-making capabilities may provide some performance advantage over other low-level automation conditions. This result is consistent with the finding of Endsley and Kaber (1999) that higher LOAs, including blended decision making, are effective in reducing task errors in individual performance with automation. There appeared to be a performance breakdown (see Figure 5) associated with decision automation under conditions of high task difficulty. Although this does not directly support our hypothesis that differences between automation conditions will be more apparent at high workload (H10), it does support the notion that potential problems may become apparent at high levels of task difficulty.

Regarding the effects of automation on team coordination in general, our findings were consistent with previous research assessing the effects of different LOAs on human performance (Endsley & Kaber, 1999; Kaber, Wright, & Clamann, 2002; Laois & Giannacourou, 1995; Ruff, Draper, & Narayanan, 2000). Automation of early and intermediate stages of human-machine system information processing may have benefits with respect to teamwork, whereas decision automation may provide benefits in more limited contexts. However, this general inference must be considered in light of other factors that may interact with the automation design in influencing team coordination. For example, the issue of workload distribution between team members may be critical to the effectiveness of automation. In the TDT, the IO was more highly loaded than was the AC. Although the information acquisition automation reduced the need for communication between team members, it did so by adding to the IO's already high visual load. It is possible that this why we did not see improvements in team coordination and effectiveness for this automation manipulation. Information analysis automation affected the workload distribution between team members in a different way. Although information analysis did not reduce task-relevant communications, it did appear to promote a strategy that reduced the IO's need for visual search and resulted in higher ratings of team skill.

CONCLUSIONS

In general, the results of this study have practical implications for automation design in team tasks to support team member coordination, including support for the use of early and intermediate forms of automation related to the acquisition and analysis of information. Although it is difficult to make specific automation design recommendations that generalize to a wide variety of team tasks, our experiment provided insight regarding the process of designing automation for team tasks. First, we recommend the use of a task-analytic approach that considers the potential effects of automation on the workload of both team members as well as on team coordination strategies. For example, there may be ways to reduce communication bottlenecks through the automation of information acquisition. If so, additional information should be provided in a manner that doesn't lead to visual or other forms of overload. Beyond this, automation of information analysis functions should be used to support efficient coordination between team members.

Second, we recommend that proposed automation design changes be evaluated through methods that include analysis of variation in communication patterns as well as team coordination, workload, and performance measures. The measurement of task-specific communications in this experiment provided insight into the way automation influenced teamwork that would not have been apparent from counts of total communications or ratings of team coordination skills alone. Evaluations should incorporate the range of task load and consider the implications of automation failures, particularly if decision automation is to be considered.

Finally, some limitations to this study should be noted. We did not systematically manipulate automation of our decision-making task to allow for comparison of each and every form of automation that may be described by the model of types and levels of automation put forth by Parasuraman et al. (2000). Rather, we considered a subset of conditions that provided some comparisons related to relative differences in certain combinations of automation expected in real-world systems. The automation manipulations, for the most part, affected the role of only the IO. Future research aimed at further explaining team performance effects of automation would include an investigation of the main effects and interactions of automation manipulations along each of the four information-processing functions identified by Parasuraman et al. (2000) and for each of the team members, together and in isolation, as recommended by Jentsch and Bowers (1996). In addition, we considered only the effects of highly reliable automation. Future research may also consider how changes in automation reliability may affect team performance.

Even though we included multiple measures of team communication and coordination, the measures depended on observation of specific communications. More direct measures of team knowledge or team SA (Cooke, Kiekel, & Bell, 2002; Endsley, 1995) may provide further insight into effects of various forms of automation on team SA and performance. Finally, although we provided some insight into the potential interactions between various forms of automation and team member workload distribution and team coordination strategies, future research is warranted to more fully explore the effects of various forms of automation on team coordination, SA, and performance.

ACKNOWLEDGMENTS

This research was supported by the National Aeronautics and Space Administration (NASA) Langley Research Center, Crew/Vehicle Integration Branch, under Grant Nos. NAG-1-01039 and NAG1-01-002. The grant monitors included Lance Prinzel and Anna Trujillo. The research was completed while the first author worked as a research associate and the second author worked as an associate professor in the Department of Industrial Engineering at North Carolina State University. Some support for the writing of this paper by the first author was provided by SA Technologies and by the Duke University Medical Center Department of Anesthesiology. The views and conclusions presented in this paper are those of the authors and do not necessarily reflect the views of NASA. We would like to thank Srikanth Jayanthi for his input on the design of the team rating scales and his efforts as a rater, Cheryl Bolstad and Mark Bolstad for their assistance with the TDT, and Mica Endsley for her review of a previous draft. We would also like to thank the anonymous reviewers for many insightful comments through their reviews of the manuscript.

REFERENCES

Bainbridge, L. (1987). Ironies of automation. In J. Rasmussen, K. Duncan, & J. Leplat (Eds.), *New technology and human error* (pp. 271–283). New York: Wiley.

Bolstad, C. A., & Endsley, M.R. (1999). Shared mental models and shared displays: An empirical evaluation of team performance. In *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting* (pp. 213–217). Santa Monica, CA: Human Factors and Ergonomics Society.

Bolstad, C. A., & Endsley, M. R. (2000). The effect of task load and shared displays on team situation awareness. In *Proceedings of the XIVth Triennial Congress of the International Ergonomics Association and 44th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 1.189–1.192). Santa Monica, CA: Human Factors and Ergonomics Society.

Bowers, C. A., Deaton, J., Oser, R., Prince C., & Kolb, M. (1993). The impact of automation on crew communication and performance. In *Proceedings of the Seventh International Symposium on Aviation Psychology* (pp. 758–761). Columbus: Ohio State University.

Bowers, C. A., Jentsch, F., & Salas, E. (1994). Studying automation in the lab – Can you? Should you? In N. McDonald, N. Johnston, & R. Fuller (Eds.), *Applications of psychology to aviation systems* (Vol. 3, pp. 281–286). Avebury, UK: Ashgate.

Bowers, C. A., Morgan, B. B., Salas, E., & Prince, C. (1993). Assessment of coordination demand for aircrew coordination and training. *Military Psychology, 5,* 95–112.

Bowers, C. A., Oser, R. L., Salas, E., & Cannon-Bowers, J. A. (1996). Team performance in automated systems. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 243–263). Mahwah, NJ: Erlbaum.

Brannick, M. T., Prince, A., Prince, C., & Salas, E. (1995). The measurement of team process. *Human Factors, 37,* 641–651.

Brannick, M. T., Roach, P. M., & Salas, E. (1993). Understanding team performance: A multimethod study. *Human Performance, 6,* 287–308.

Clamann, M. P., Wright, M. C., & Kaber, D. B. (2002). Comparison of performance effects of adaptive automation applied to various stages of human-machine system information processing. In *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting* (pp. 342–346). Santa Monica, CA: Human Factors and Ergonomics Society.

Clothier, C. (1991). Behavioral interactions across various aircraft types: Results of systematic observations of line operations and simulations. In *Proceedings of the Sixth International Symposium on Aviation Psychology* (pp. 332–337). Columbus: Ohio State University.

Cooke, N., Kiekel, P. A., & Bell, B. (2002). Addressing limitations in the measurement of team cognition. In *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting* (pp. 403–407). Santa Monica, CA: Human Factors and Ergonomics Society.

Costley, J., Johnson, D., & Lawson, D. (1989). A comparison of cockpit communication B737-B757. In *Proceedings of the Fifth International Symposium on Aviation Psychology* (pp. 413–418). Columbus: Ohio State University.

Coury, B. G., & Semmel R. D. (1996). Supervisory control and the design of intelligent user interfaces. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 201–219). Mahwah, NJ: Erlbaum.

Dickinson, T. L., & McIntyre, R. M. (1997). A conceptual framework for teamwork measurement. In M. T. Brannick, E. Salas, & C. Prince (Eds.), *Team performance assessment and measurement: Theory, methods and applications* (pp. 19–45). Mahwah, NJ: Erlbaum.

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

Endsley, M. R. (1997). Level of automation: Integrating humans and automated systems. In *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting* (pp. 200–204). Santa Monica, CA: Human Factors and Ergonomics Society.

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

Foushee, H. C., Lauber, J. K., Baetge, M. M., & Acomb, D. B. (1986). *Crew factors in flight operations III: The operational significance of exposure to short-haul air transport operations* (NASA Tech. Memorandum 88322). Moffett Field, CA: NASA-Ames Research Center.

Galster, S. M., Bolia, R. S., & Parasuraman, R. (2002). Effects of information automation and decision-aiding cuing on action implementation in a visual search task. In *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting* (pp. 438–442). Santa Monica, CA: Human Factors and Ergonomics Society.

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. 139–178). Amsterdam: Elsevier Science.

Jentsch, F., & Bowers, C. (1996). Automation and crew performance: The importance of who and what. In *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting* (pp. 49–53). Santa Monica, CA: Human Factors and Ergonomics Society.

Jentsch, F., Sellin-Wolters, S., Bowers, C. A., & Salas, E. (1995). Crew coordination behaviors as predictors of problem detection and decision making times. In *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting* (pp. 1350–1353). Santa Monica, CA: Human Factors and Ergonomics Society.

Johannesen, L. J., Cook, R. I., & Woods, D. D. (1994). Cooperative communications in dynamic fault management. In *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting* (pp. 225–229). Santa Monica, CA: Human Factors and Ergonomics Society.

Kaber, D. B., & Endsley, M. R. (1997). The combined effect of level of automation and adaptive automation on human performance with complex, dynamic control systems. In *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting* (pp. 205–209). Santa Monica, CA: Human Factors and Ergonomics Society.

Kaber, D. B., Onal, E., & Endsley, M. R. (2000). Design of automation for telerobots and the effect on performance, operator situation awareness and subjective workload. *Human Factors and Ergonomics in Manufacturing, 10,* 409–430.

Kaber, D. B., & Riley, J. R. (1999). Adaptive automation of a dynamic control task based on secondary task workload measurement. *International Journal of Cognitive Ergonomics, 3,* 169–187.

Kaber, D. B., Wright, M. C., & Clamann, M. P. (2002). *Human response to adaptive automation of early sensory/information acquisition functions and later stages of information processing in complex system operations* (Final Report, NASA Langley Research Center Grant #NAG-1-01039). Hampton, VA: NASA Langley Research Center.

Laois, L., & Giannacourou, M. (1995). Perceived effects of advanced ATC functions on human activities: Results of a survey on controllers and experts. In *Proceedings of the Eighth International Symposium on Aviation Psychology* (pp. 392–397). Columbus: Ohio State University.

MacMillan, J., Entin, E., & Serfaty, D. (2002). From team structure to team performance: A framework. In *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting* (pp. 408–412). Santa Monica, CA: Human Factors and Ergonomics Society.

Orasanu, J. M. (1990). *Shared mental models and crew decision making* (Cognitive Science Laboratory Report No. 46). Princeton, NJ: Princeton University.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, 30,* 286–297.

Rovira, E., Zinni, M., & Parasuraman, R. (2002). Effects of information and decision automation on multi-task performance. In *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting* (pp. 327–331). Santa Monica, CA: Human Factors and Ergonomics Society.

Ruff, H., Draper, M. H., & Narayanan, S. (2000). The effect of automation level and decision aid fidelity on the control of multiple remotely operated vehicles. In D. B. Kaber & M. R. Endsley (Eds.), *Human performance, situation awareness and automation: User-centered design for the new millennium* (pp. 70–75). Madison, WI: Omnipress.

Sarter, N. B., & Schroeder, B. (2001). Supporting decision making and action selection under time pressure and uncertainty: The case of in-flight icing. *Human Factors, 43,* 573–583.

Scerbo, M. W. (1996). Theoretical perspectives on adaptive automation. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 37–63). Mahwah, NJ: Erlbaum.

Serfaty, D., Entin, E. E., & Johnston, J. (1998). Team coordination training. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress* (pp. 221–245). Washington, DC: American Psychological Association.

Sheridan, T. B., & Verplanck, W. L. (1978). *Human and computer control of undersea teleoperators* (Tech. Report). Cambridge, MA: MIT Man-Machine Laboratory.

Travillian, K. K., Volpe, C. E., Cannon-Bowers, J. A., & Salas, E. (1993). Cross-training highly interdependent teams: Effects on team processes and team performance. In *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting* (pp. 1243–1247). Santa Monica, CA: Human Factors and Ergonomics Society.

Urban, J. M., Bowers, C. A., Monday, S. D., & Morgan, B. B. (1993). Effects of workload on communication processes in decision-making teams: An empirical study with implications for training. In *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting* (pp. 1233–1237). Santa Monica, CA: Human Factors and Ergonomics Society.

Volpe, C. E., Cannon-Bowers, J. A., Salas, E., & Spector, P. E. (1996). The impact of cross-training on team functioning: An empirical investigation. *Human Factors, 38,* 87–100.

Wickens, C. D., Mavor, A. S., Parasuraman, R., & McGee, J. P. (Eds.). (1998). *The future of air traffic control: Human operators and automation.* Washington, DC: National Academy.

Wiener, E. L. (1993). Crew coordination and training in the advanced-technology cockpit. In E. L. Wiener, B. G. Kanki, & R. L. Helmreich (Eds.), *Cockpit resource management* (pp. 199–266). San Diego, CA: Academic.

Wise, J. A., Guide, P. C., Abbott, D. W., & Ryan, L. (1992). *Human factors in aviation safety: The effects of automation on corporate pilots* (Interim Tech. Report CAAR-15406-92-1). Daytona Beach, FL: Embry-Riddle Aeronautical University.

Woods, D. D. (1996). Decomposing automation: Apparent simplicity, real complexity. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 3-16). Mahwah, NJ: Erlbaum.

Melanie C. Wright is an assistant professor in the Department of Anesthesiology and the Human Simulation and Patient Safety Center at Duke University Medical Center. She received her Ph.D. in industrial engineering at North Carolina State University in 2002.

David B. Kaber is a visiting professor at the University of Rostock (Germany) and an associate professor in the Department of Industrial Engineering at North Carolina State University. He received his Ph.D. in industrial engineering at Texas Tech University in 1996.

Date received: October 17, 2002
Date accepted: January 18, 2005