Evaluation of Semi-Autonomous Convoy Driving

James Davis and Asisat Animashaun
U.S. Army Research Laboratory
Aberdeen Proving Grounds, Maryland 21001
e-mail: james.davis@arl.army.mil,
aanimashaun@arl.army.mil

Edward Schoenherr
U.S. Army Tank and Automotive Research Development and Engineering Center
6501 E. Eleven Mile Road
Warren, Michigan 48397
e-mail: edward.schoenherr@us.army.mil

Kaleb McDowell
U.S. Army Research Laboratory
Aberdeen Proving Grounds, Maryland 21001
e-mail: kmcdowell@arl.army.mil

Received 9 April 2008; accepted 27 August 2008

Autonomous mobility technologies may have applications to manned vehicle convoy operations—they have the ability to enhance both system performance and operator capability. This effort examines the potential impact of introducing semi-autonomous mobility [Convoy Active Safety Technologies (CAST)] into manned vehicles. Twelve civilians with experience driving military vehicles in convoy-type operations participated in this experiment. For the experiment, they were tasked with following a lead vehicle while completing a concurrent security task (scanning the local environment for targets). The control of the manned vehicle was varied between CAST and manual control at several different speed levels. Several objective speed and accuracy variables along with subjective operator assessment variables were examined for each task. The results support the potential benefits of incorporating semi-autonomous mobility technologies into manned vehicle convoy operations. The semi-autonomous mobility system was associated with significantly better performance in several aspects of operator situational awareness and convoy integrity, including enhanced target identification, improved maintenance of following distance, and improved performance for unanticipated stops. This experiment also highlighted a critical human factors issue associated with the incorporation of autonomy in real-world applications: participants felt that, overall, they...
Sustainment convoys are critical to providing combatant commanders the right support, at the right time and place, and in the right quantities, across the full range of military operations. The ability to conduct sustainment convoys in a variety of hostile environments requires force protection measures that address the enemy threat and protect the soldier. This issue directly impacts the vehicle driver, in that the security issue adds such high demands on operators that it fundamentally changes the driving task from that of the civilian driving to that of secure mobility—secure mobility is defined as the capability of the soldier-system to transverse terrain in a manner that meets mission demands (vehicle mobility) while maintaining a real-time understanding of the environment local to one’s vehicle and platoon (local area awareness) (Nunez, 2006). A potential solution to secure mobility is to leverage autonomous mobility technologies created for robotic assets for the purpose of reducing mobility demands on vehicle operators and increasing local area awareness capabilities. Robotic follower technologies offer one potential solution and could be transitioned into a semi-autonomous convoy capability much in the same way that lane maintenance technologies have been proposed for transition to civilian vehicles (Rossetter, Switkes, & Gerdes, 2004). This investigation focuses on the potential impact of semi-autonomous mobility technologies on overall convoy performance and operator local area awareness, workload, and fatigue.

1.1. Background

In civilian settings vehicle design is geared toward safety rather than security: the risks that civilian drivers face do not come from hostile agents. Civilian vehicle safety requirements generally relate to providing an operator with the ability to see and avoid cars, pedestrians, and other obstacles around the vehicle as the vehicle traverses well-defined, marked roadways and performs maneuvers such as parking and passing. The operational execution of the secure mobility function is conceptually different from that of civilian driving. In the military setting, the early detection of hostile threats becomes a critical and demanding component of the overall driving. The security demands in addition to the traditional civilian driving tasks create an environment where soldiers must perform multiple high-workload tasks concurrently, a situation that multiple resource theory suggests can lead to performance decrements due to the nonavailability of “workload resources” (Wickens, 1980).

The military’s initiative to reduce manpower has limited the solution space for reducing the task loading associated with driving a military convoy. Designing more intuitive user interfaces for the convoy driver affords some reductions in task loading by reducing the effort associated with completion of the task. Through manpower, personnel, and training (MPT) analyses the optimal allocation of tasks among the crew members can be determined. However, with current military manning objectives, research has shown that even with the optimal task allocation, task loading is still very high (Mitchell, Samms, Henthorn, & Wojciechowski, 2003). Aside from designing more intuitive driver interfaces and conducting MPT analyses, automation is one of the few remaining solutions. Autonomous mobility technologies have the potential to improve overall convoy performance by reducing specific vehicle control demands on the vehicle operators (Young & Stanton, 2004), improving vehicle control, reducing training time, preventing collisions, and improving vehicle safety (Ackermann & Bunte, 1997; Kasselmann & Keranen, 1969; Stanton & Young, 1998; Yih & Gerdes, 2005). Specific applications of autonomous mobility technologies to manned convoy vehicles include but are not limited to providing route information to the operator, active steering, overriding control inputs to “safeguard” against accidents, and virtual drivers (Mammar, Sainte-Marie, & Glaser, 2001), which may perform a variety of tasks including lane maintenance or full-on path selection. Adjustable or “sliding” autonomy such as collaborative control offers potentially the greatest impact on manned convoy operations because it
allows for exploitation of desirable human qualities (e.g., mission planning and local security) without sacrificing the benefit of automation (e.g., basic path following and collision avoidance) (Crandall, Goodrich, Olsen, & Nielsen, 2005).

The goals of incorporating autonomous motility technologies are similar to those associated with the incorporation of autonomous motility technologies such as adaptive cruise control into the civilian arena: safely increasing traffic throughput (Bengtsson, 2001; Ioannou & Chien, 1993; Vahidi & Eskandarian, 2003). However, the additional task loading associated with mobile security necessitates an additional goal of reducing the driver’s task loading. Successful implementation of such autonomy has enhanced vehicle mobility and reduced operator workload for a variety of unmanned vehicles in specific environments (e.g., space exploration and industrial security). However, previous work demonstrates that autonomous technologies do not always increase overall performance (Endsley & Kaber, 1999; Sheridan, 1992; Young & Stanton, 1997). Integrating autonomous mobility technologies will not simply remove tasking from an operator, but rather will change the responsibilities of the operator (Sheridan, 1992) and consequently the responsibilities of the entire crew within the vehicle. Shifts in task allocations generally have both benefits and costs (Sheridan & Parasuraman, 2000) and are dependent on fluid human factors such as operator trust in the automation, mental workload, cognitive overhead, confidence, and risk (Parasuraman & Riley, 1997). If implemented effectively, the task allocations within the overall crew-automation team will be restructured in a manner that reduces the overall demands on the operator (and crew) and potentially enhances performance directly through increased convoy integrity or indirectly through related factors such as reduced fatigue. Whereas the promise of autonomous mobility technologies is to potentially reduce demands on the driver and increase overall crew-system performance, it is important to verify these performance enhancements in an operationally relevant environment (paved as well as off-road terrain following) under operationally relevant stressors (i.e., while the operator is completing local security tasks).

1.2. Purpose

This effort examines human interaction with semi-autonomous systems for the purpose of enhancing secure mobility performance during convoy operations. Previous efforts have examined the successful incorporation of autonomy in nearly all aspects of civilian driving including speed maintenance (Tanigawa, Masuda, Nakamura, & Hayashi, 1984), steering (Franke, Mehring, Suissa, & Hahn, 1994), lane changing (Jochem, Pomerleau, & Thorpe, 1995), parking (Xu, Chen, & Xie, 2000), highway driving (Dickmanns & Zapp, 1985), and urban driving (Franke et al., 1998). However, this experiment examined the incorporation of a semi-autonomous system in an operational environment similar to that experienced on military convoys. McDowell, Nunez, Hutchins, and Metcalfe (2008) completed a similar experiment that incorporated a semi-autonomous system into a military convoy. McDowell et al. (2008) demonstrated the ability of a semi-autonomous system to reduce workload (i.e., effort and physical demand) and enhance local security (i.e., target detection reaction time); however, they examined very few measures associated with maintaining convoy integrity. In addition, the McDowell et al. (2008) study examined the incorporation of a semi-autonomous system into proposed military vehicles of the future (i.e., drive-by-wire systems), whereas this experiment examined the incorporation of a semi-autonomous system into current military convoys (i.e., direct vision systems). For this experiment, the operators were responsible for concurrently maintaining a designated following distance (FD) [as determined by the Army’s Field Manual for Wheeled Vehicles (FM 21-305)], responding to unanticipated stops by the lead vehicle (LV), and scanning the local environment for targets. Unanticipated stops were incorporated to examine the ability of the participant to react to abrupt changes in the vehicle convoy while completing the local security task—with and without the autonomous system. Driving and scanning tasks were completed while traversing both paved and unpaved roads at speeds ranging from 20 to 60 km/h. Two control modes were examined: (1) manual driving and (2) semi-autonomous mobility using the Convoy Active Safety Technologies (CAST) system (Axe, 2008). During the semi-autonomous control mode, participants took a supervisory role (Sheridan, 1992) when controlling the vehicle, intervening only if the vehicle veered off of the path. Several objective and

1The CAST system has been developed as an affordable semi-autonomous technology alternative to assist in convoy operations.
subjective measures were examined for each control mode. It was hypothesized that incorporation of the semi-autonomous driving system into the experimental convoy operation would result in enhanced convoy integrity (i.e., more consistent FDs, shorter braking distances, and less deviation from the track of the LV) and reduced driver task loading (as determined by their performance on a secondary security/scanning task and objective measures such as workload and situational awareness).

2. METHOD

2.1. Participants
Owing to the limited subject pool of people with experience driving 2.5- to 5-ton military vehicles in convoy-type operations, participants were contracted through TechWise (Ft. Eustis, Norfolk, Virginia), a company that specializes in providing drivers of military vehicles. The voluntary and fully informed consent of all of the participants in this experiment was obtained prior to participation in this study as required by CFR 219 and AR70-25. Twelve civilians (eight male, four female) participated in this experiment. Of the 12 participants, 8 had prior military service experience. All participants had current and valid commercial driver’s licenses for 2.5- to 5-ton vehicles and had prior experience driving heavy military vehicles in convoy-type operation. The average age of the participants was 43.25 ± 9.4 years. This age group was not optimal; as of 2005 only 32% of enlisted soldiers were over the age of 30 (Maxfield, 2005); however, it was tolerated due to the limited availability of people with convoy driving experience.

2.2. Apparatus

Test vehicles: Two Family of Medium Tactical Vehicles (FMTVs) A1 wheeled trucks were used in this experiment (Figure 1). The two FMTVs represented a LV and follower vehicle (FV) in the simulated convoy scenario. Both the LV and FV were retrofitted with the CAST system for semi-autonomous driving. An experimenter operated the LV, and the participant operated the FV.

CAST system: The CAST system is a semi-autonomous technology that utilizes relative positioning to follow a LV (maximum FD is 100 m) along a discrete path while avoiding collisions with the LV (Axe, 2008). The CAST system is essentially a vehicle follower with path following capabilities. The FV uses a perception vehicle follower (PVF) system to complete vehicle following. The PVF estimates the relative position of the LV with respect to a coordinate system centered on the FV. The relative positions of the LV positions form a track that may be followed by the PVF system’s guidance and control subsystems. PVF cycles at roughly 10 Hz, with each cycle consisting of the following steps: (1) read new data from sensors, (2) estimate LV position, and (3) calculate steering and speed command. The PVF leverages data from multiple sensors so that it is robust to dropouts or failures in one or more sensors. The primary sensors used by the PVF system include a radar system and a color camera system. A rough measurement of LV position and relative speed is first determined using the radar system. Then processing of the area of the camera image (specified by the radar location) is used to determine a more accurate measurement of LV position and speed. If the radar system fails, image processing from the color camera will

Figure 1. Apparatus. (A) FMTV. (B) Convoy showing lead and follower vehicles.
still be carried out, only over a larger image area and with consequently longer processing times. If the image processing from the color camera fails, then the PVF will default to less-accurate speed and position measurements based only on the radar system. The radar system is accurate enough to support rudimentary vehicle following; however, the resulting cross-track error will often be impractical for use on narrow roads.

In cases when the LV’s PVF system is unable to track the LV (e.g., when the LV takes a really sharp corner or there is limited visibility due to dust), the real-time path following allows the FV to traverse the environment. Real-time path following is achieved through the use of the color camera system, a SICK LADAR system, and a global positioning system (GPS). The first step in path following is identifying the location of the road. It is assumed that the road is a fairly smooth surface when compared with grass and bushes that surround the road. By comparing the elevations of the road determined from the SICK LADAR system, the nominal location of the road can be determined. The nominal location of the road can then be projected on the color camera image, and the road location can be verified based on the color at that pixel. Unfortunately, there are situations in which the SICK LADAR system does not produce accurate data; e.g., if the road is wet many of the laser beams will reflect off the road and never return, causing the SICK LADAR system to report very large ranges. When the FV is unable to collect accurate data from the SICK LADAR system, it ignores the inaccurate data and checks back every few seconds to see whether the problem has resolved itself. However, as long as the SICK data are unreliable, image processing from the color camera will still be carried out (i.e., searching for pixels to match the colors of the road), only over a larger processing area and with consequently longer processing times. If the LADAR and color camera systems fail or the LV is out of the range of the radar and color camera systems, the LV will default to GPS data. If a degradation in the systems’ sensing abilities (whether in vehicle follower or path follower mode) leads to large deviations from the desired course of travel, the operator could simply disengage the CAST system, realign the vehicle behind the LV, and then reengage the CAST system (see section Operator Station below for more details).

Steering and speed control were maintained differently depending on whether the CAST system was in vehicle following or path following mode. When in vehicle following mode (i.e., when the CAST system could locate the LV), steering and speed were calculated based on the radar data from the PVF system and a predefined FD. However, when the LV completes a sharp turn, radar data will cause the vehicle to cut the turn short because they attempt to steer straight toward the LV. So when the LV turns, the CAST system defaults to path following mode. When in path following mode (i.e., when the CAST system could not locate the LV), steering was calculated solely using the SICK LADAR information and the GPS trail from the LV and speed was matched to that of the LV (to avoid rear-end collisions). Unfortunately, the GPS was far less accurate than the radar system and sometimes led the vehicle to steer away from the designated path. Fortunately, GPS navigation is used very rarely during CAST vehicle control (less than 1% of the time) due to its inaccuracy and due to the fact that it is often being jammed by our forces or enemy forces in operational environments (Magnuson, 2008).

Operator station: The operator station consisted of the steering controls, a disengage pedal, a confidence display, and a target button. The steering controls included the steering wheel and brake and gas pedals. Steering controls were used when the system was not in semi-autonomous mode or when the participant disengaged the semi-autonomous system. The disengage pedal was located to left of the gas pedal in a location similar to that of the clutch on a vehicle with manual transmission [see Figure 2(A)]. The disengage pedal allowed the participant to disengage the CAST system. The confidence display was located just below the windshield [see Figure 2(B)]. The confidence display provided the participant with information regarding the confidence of the system (i.e., sensor’s ability to identify the LV and road), the FD, and the status of the system (e.g., engaged or disengaged). Also located on the confidence display was the button (with green LED) that allowed the participant to put the system in semi-autonomous control mode at the beginning of a trial or after the participant pressed the disengage pedal. Located above the confidence display was a (yellow and red) emergency disengage button that the experimenter could press if he or she needed to disengage the system. The target push button was located on the steering wheel well and was used by the participant to indicate when a target was detected [see Figure 2(C)]. Input from the steering controls, disengage button, confidence display, and target push button and
Figure 2. Apparatus. (A) Pedal used to disengage semi-autonomous CAST system. (B) CAST confidence display that provided information regarding the status of the CAST system. (C) Target push button used for the scanning task.

disengage pedal were integrated into the CAST system so that it could be synchronized with position and time information. Time-synced data from the operator station and position data were used to determine the participant’s response time for target identification and unanticipated stopping of the LV.

Driving course: The test course was approximately 19 km in length and was generally flat with several hills and turns; see Figure 3.

Cognitive test: Between each driving trial, the participants completed a reaction-time and short-term-memory test from the computer-based Automated Performance Test System (APTS) (RSK Associates, Inc., Orlando, Florida). For the reaction-time test, four empty boxes appeared on the screen and participants were responsible for pressing the number key corresponding to the illuminated box. For the short-term-memory test, participants were presented with a group of four letters. After memorizing these letters, participants were presented consecutively with 10 letters and tasked with identifying whether each letter was a part of the initial group of four letters by using the “S” (same) and “D” (different) buttons on the keyboard.

Figure 3. Test course with actual test route highlighted. The upper portion of course was unpaved (i.e., dirt), and the lower portion of the course was paved.
Subjective measures: Four subjective questionnaires were administered between each driving trial. A modified computerized version of the NASA Task Loading Index (NASA-TLX) (Hart, Muth, Mordkoff, Levine, & Stern, 2001) was used to collect workload data that assessed workload along the six subscales. A computerized version of the Situational Awareness Rating Technique (SART) Questionnaire (Taylor, 1990) was used to assess situational awareness between the driving trials. A computerized version of the Motion Sickness Assessment Questionnaire (MSAQ) (Gianaros, Muth, Mordkoff, Levine, & Stern, 2001) provided a subjective measure of participant motion sickness. A trust-rating scale was used to assess the participant’s level of trust in the automated driving system (Jian, Bisantz, & Drury, 2000). The trust-rating scale was administered only between driving trials that incorporated the CAST system. In addition to these four subjective measures administered between the driving trials, a basic fatigue-rating scale (“Please rate your level of fatigue on a scale from 1–10, where 1 is ‘wide awake’ and 10 is ‘dead tired’”) was used to assess the participants’ fatigue levels before and during each driving trial. After the experiment, an exit interview was conducted to assess several issues regarding the usability of the CAST system. Prior to starting the experiment a demographics questionnaire was also completed.

2.3. Procedures

This experiment took place over 2 days. The first day was designated for general orientation to the CAST system and test course, and the second day was the actual test day. Upon arrival on the first day (orientation day), participants provided their voluntary consent to participate in the experiment and completed a demographics questionnaire. During orientation, participants received both classroom and hands-on experience. They were trained on and practiced how to operate the vehicle with the semi-autonomous system engaged. The designated FD was demonstrated over the 5-km stretch as the participant followed the LV to the test course. Participants were not provided any training on operation of the FMTV, because all participants operated FMTVs (or similar vehicles) in convoy-type operations on a regular basis as required by their occupations. Participants were also trained on how to complete all questionnaires and assessments.

On the second day (within 2 weeks of the orientation), participants completed the testing segment of the experiment. Upon arrival at the base station, they completed a set of assessments to determine their baseline levels of cognition, motion sickness, and fatigue.

Participants then completed seven trials during this experiment: one refresher trial and six experimental trials. The first driving trial was a refresher trial on the CAST system. It consisted of the participants traversing the test course twice with the CAST system on. The participants then completed a set of six driving trials: three with the CAST system engaged and three with the CAST system disengaged. Each driving trial was traversed at a designated speed level (low, medium, and high) set by the LV. For the low-speed level, participants traveled 20 km/h over the unpaved terrain and 40 km/h over the paved terrain. For the medium-speed level, participants traveled 30 km/h over the unpaved terrain and 50 km/h over the paved terrain. For the high-speed level, participants traveled 40 km/h over the unpaved terrain and 60 km/h over the paved terrain. Driving speed level and control mode (manual/semi-autonomous) were counterbalanced to account for order effects. The total experiment took approximately 7 h and 45 min, including breaks and lunch.

During each driving trial, the participant had three responsibilities: scanning the environment for targets (20 per trial), maintaining a designated FD (40 m), and reacting to unanticipated stopping of the LV (four per trial). Targets included large orange construction barrels, brown caution cones, large green cones, black trash cans, and small yellow cones. Participants reported the detection of a target by depressing the target push button located on the steering wheel well and verbally describing the general location and type of target detected (e.g., "orange barrel, right"). The prescribed FD was determined according to the Field Manual for Wheeled Vehicles (FM 21-305) for safe driving distances at the maximum speed in this experiment (60 km/h). For unanticipated stops, participants were instructed to stop the vehicle as quickly as possible without skidding. When the CAST system was engaged, it automatically responded to stopping of the LV. Participants experienced a total of four unanticipated stops per driving trial: two on each type of terrain. Unanticipated stops occurred only on straightaways and were counterbalanced to account for order effects.
When the CAST system was engaged, participants were responsible for monitoring the position of their vehicle within the environment. If the vehicle strayed off course or encroached upon the designated FD, participants were responsible for disengaging the CAST system and realigning the vehicle. Once the vehicle was realigned, the participant could then reengage the system. Two experimenters were in the FV with the participant at all times to monitor the participant and the CAST system. The participant or experimenter in the FV could disengage the CAST system using the disengage button or the emergency disengage button, respectively. After disengaging the system, the CAST system had to be reengaged using the button on the CAST confidence display in order revert back to the semi-autonomous control mode.

2.4. Data Reduction

2.4.1. Convoy Integrity

Measures of convoy integrity were derived from the CAST system and include the following.

Maintenance of FD: FD was determined by calculating the distance between coordinates of both vehicles. Both the mean and standard deviation (STD) of the FD were computed.

Reaction to unanticipated stops—Braking distance (BD) and braking reaction time (BRT): As participants were instructed to stop as quickly as possible when the LV stopped, the BD was used to examine a participant’s ability to react to unanticipated stops. BD was defined as the distance it took the FV to come to a full stop following the initiation of braking by the LV. BD was calculated as the distance between the position of the FV at the time when the LV first applies the brakes and the time when FV first comes to a full stop and was averaged across the unanticipated stops within a control mode and terrain type. BRT was used as a supplemental measure to analyze reaction to unanticipated stopping of the LV. BRT was the difference in time between the driver in the LV pressing the brake pedal and the operator or semi-autonomous system in the FV pressing the brake pedal.

Lateral offset (LO): LO was calculated as the lateral distance between vehicle paths as measured from the centerline of the each vehicle. Both the mean and STD of the LO were computed.

Course corrections (CC), E-stop, and system failures: Given the maturity of the technology, the system sometimes suffered from impaired sensory information (i.e., the system would have a difficult time locating the LV). This would result in the vehicle deviating slightly from the centerline of the path defined by the LV. When this was significant enough, the participant would apply a CC. That is, the operator would disengage the system, realign the vehicle to the LV ahead, and then reengage the system. The total number of CCs per speed level was calculated per driving trial. In cases when the operator did not apply a CC in a reasonable amount of time, the experimenter administered an E-stop. This would disengage the semi-autonomous control and allow the participant to take over manual control of the vehicle. The total number of E-stops was calculated per driving trial. The duration and frequency of system failures (i.e., times when the system had to be stopped and repaired) were also recorded. The frequency and duration were used to determine the average duration of a system failure.

Overall mission completion time: Overall mission time was measured by the amount of time it took the convoy to traverse the course from start to finish. Training trials and system failures were not included in overall mission times.

2.4.2. Local Area Awareness

Participants’ performance on the local area awareness (target identification) task was reduced as calculated in terms of speed and accuracy.

Response time (RT): The speed with which participants were able to identify the target was determined by using vehicle position information. A GPS location just beyond the point of first target recognition was determined for each target—a ‘line of sight’ location. The speed (RT) with which the participant was able to identify the target was calculated by subtracting the time at which the participant passed the line of sight location from the time at which the participant depressed the target button. To account for differences in the type and placement of targets in this natural environment, which are known to cause differences in response time, RTs were normalized across each of 60 unique target placements (20 targets × 3 target schemes) using a z score.

Percent correct (%C): The percentage of targets correctly detected (%C) was determined via verbal report by the participant.
2.4.3. Posttrial Operator Assessments

Several assessments were conducted after each driving trial, including the following:

- **Fatigue:** Data from the physiological measurement equipment will not be reported in this paper. However, this paper will provide a preliminary look at fatigue from the subjective responses on the fatigue-rating scale.
- **Workload:** The NASA-TLX workload scale was reduced along six subscales: mental demand, physical demand, temporal demand, effort, frustration, and performance.
- **Overall situational awareness:** The 10-question SART was examined across the three major dimensions: attentional demand, attentional supply, and understanding.
- **Motion sickness:** The MSAQ was reduced along four components of motion sickness: gastrointestinal, central, peripheral, and sopite-related measures.
- **Cognition:** The percent of correct answers and reaction time were examined for both the reaction-time and short-term-memory tasks.
- **Trust:** Trust in the CAST system was assessed using the 10-question Trust Survey (CAST), which will be superseded by the interaction and is not discussed.

2.4.4. Exit Interview

Results from an exit interview on system performance were summarized.

2.5. Statistics

Linear mixed models (McCulloch, 2003) were used to analyze the effects of the CAST system across the different speed levels and control modes. The covariance structure was variance components for all linear mixed models. Post hoc evaluations were pairwise comparisons using the Bonferroni method. Linear mixed models included speed level (low, medium, high) and control (CAST engaged, manual/CAST disengaged) mode as fixed factors and participant as a random effect. For the trust survey and CC measure, the linear mixed model included speed level only as a fixed factor (and participant as a random factor) because it was administered/measured for only one control mode (i.e., semi-autonomous control). All statistics were reported with outliers removed. Outliers and missing data accounted for less than 1% of the data and were generally associated with errors. All means were reported as mean ± standard error. In addition to significant main effects, statistical trends (0.05 < p ≤ 0.10) are also reported, as these trends may become main effects with larger sample sizes.

3. RESULTS

3.1. Convoy Integrity

3.1.1. Maintenance of FD

For mean FD, a linear mixed model revealed an interaction between control mode and speed level ($F_{2,188} = 9.95, df = 187.606, p \approx 0.001$) and main effects for speed level ($F_{2,188} = 8.18, df = 187.250, p \approx 0.001$) and control mode ($F_{1,187} = 18.84, df = 187.102, p \approx 0.001$). Whereas the mean FD without CAST system engaged increased significantly with increasing speed level, the mean FD with the CAST system engaged remained the same among the different speed levels (Figure 4(A)). The main effect was superseded by the interaction and is not discussed.

For STD FD, a linear mixed model revealed an interaction between control mode and speed level ($F_{2,181} = 13.84, df = 181.341, p \approx 0.001$) and main effects for speed level ($F_{2,181} = 31.55, df = 180.999, p \approx 0.001$) and control mode ($F_{1,181} = 276.42, df = 180.825, p \approx 0.001$). Similar to the mean FD finding, pairwise comparisons revealed that when driven manually, STD FD increased significantly with increasing speed level (Figure 4(B)). The STD FD main effects for control mode and speed level were superseded by the interaction and are not discussed.

3.1.2. Reaction to Unanticipated Stops: BD and BRT

A linear mixed model for BD revealed main effects for control mode ($F_{1,224} = 58.16, df = 223.950, p \approx 0.001$) and speed level ($F_{2,223} = 22.87, df = 22.868, p \approx 0.001$); no interactions were observed. Pairwise comparisons revealed that BD with the CAST system engaged was significantly shorter than BD with the system disengaged. Pairwise comparisons also revealed that BD for low speed level was significantly shorter than BD for medium or high speed level. See Figure 5.

A mixed linear model for BRT revealed main effects for control mode ($F_{1,216} = 33.47, df = 215.611,$
Figure 4. FD for control mode by speed level interactions. (A) Average FD and (B) STD FD.

Figure 5. BD by (A) control mode and (B) speed level.

$p \approx 0.001$) and speed ($F_{2,20k} = 5.65, df = 206.125, p \approx 0.004$). Pairwise comparisons revealed nearly the same trend as BD among the different control modes; BRT with the CAST system engaged was significantly shorter than BRT with the system disengaged (see Figure 6). However, BRT among the different speed levels followed a trend opposite to that of BD. Whereas BD generally increased with increasing speed, BRT generally decreased with increasing speed. Pairwise comparisons revealed that BRT at the high speed level was significantly shorter than BRT at the low speed level (see Figure 6).

3.1.3. Lateral Offset

A main effect for control mode revealed that participants had significantly greater average LO ($F_{1,188} = 92.85, df = 188.264, p \approx 0.001$) with system engaged (0.81 ± 0.02 m) than with the system disengaged (0.52 ± 0.02 m). For STD LO, a main effect for control mode was also revealed ($F_{1,28} = 77.27, df = 27.579, p \approx 0.001$). Participants also experienced a smaller STD LO when the CAST system was disengaged (0.40 ± 0.02 m) than when it was engaged (0.60 ± 0.01). Neither an interaction nor a main effect for speed level was observed for LO or STD LO.
3.1.4. Course Corrections, E-Stops, and System Failures
For CC, there was no speed level main effect. The average CC across speed levels was 5.6 ± 0.5 or about 1 CC every 6.5 min of convoy. Per experimental observation, a CC (i.e., the participant disengaged, re-aligned the vehicle, and then reengaged the system) took less than 5 s on average. The experimenters had to press the E-stop only once, and this occurred almost simultaneously with the participant disengaging the system himself. About 58% (7 of 12) experienced a system failure. The duration for a system failure ranged from 4 to 22 min, and the average duration was 10.6 ± 3.0 min. Over the ~61 h of operation by the participants, the system was nonoperational for only about 1.5 h.

3.2. Local Area Awareness
3.2.1. Response time
A mixed linear model revealed a main effect for speed level ($F_{2,55} = 13.19, df = 55, p \approx 0.000$). RT$^2$ at the low speed level (0.31 ± 0.73) was significantly faster than RT at the medium (−0.20 ± 0.10) and high (−0.15 ± 0.09) speed levels. The model also revealed a non-significant trend between the control modes ($F_{1,55} = 3.15, df = 55, p \approx 0.082$), which suggested that participants detected objects faster with the CAST system engaged (0.07 ± 0.08) than with the CAST system disengaged (−0.09 ± 0.08).

3.2.2. Percent Correct
A mixed linear model revealed a significant main effect for control mode ($F_{1,55} = 10.02, df = 55, p \approx 0.003$), indicating that participants saw significantly more targets with the system engaged than with the system disengaged [see Figure 7(A)]. The mixed linear model also revealed a significant main effect for speed level ($F_{2,55} = 9.70, df = 55, p \approx 0.000$).

---

$^2$Reaction times are reported as normalized z scores. Higher normalized z scores equate to faster RT.
Pairwise comparisons showed that participants identified more targets at the low speed level than at the medium and high speed levels [see Figure 7(B)].

3.3. Posttrial Operator Assessments

3.3.1. Fatigue

A mixed linear model revealed no significant speed level or control mode (or interaction) main effects for self-reported fatigue. On a scale from 1 to 10 (where 1 is “wide awake” and 10 is “dead tired”), participants on average reported a fatigue score of 2.4 ± 0.2.

3.3.2. Workload

A mixed linear model revealed a significant difference of subjective reports of mental demand \( F_{1,48} = 4.04, df = 47.663, p \approx 0.050 \), performance \( F_{1,44} = 4.33, df = 44.478, p \approx 0.043 \), and effort \( F_{1,45} = 9.88, df = 44.637, p \approx 0.004 \) among the two control modes. Participants felt that they performed significantly better on the convoy operation with the CAST system disengaged (24.8 ± 5.8) than with the CAST system engaged (38.1 ± 6.2). Participants also reported that the convoy operation required significantly less effort and mental demand with the CAST system engaged (35.0 ± 5.3 and 24.9 ± 2.8, respectively) than with the CAST system disengaged (42.8 ± 6.2 and 34.5 ± 4.9, respectively). Reports of physical demand (26.9 ± 2.6), temporal demand (30.4 ± 2.9), and frustration (13.4 ± 2.8) subscales were relatively low and were not significantly different among the different speed levels or control modes.

3.3.3. Overall Situational Awareness

No significant differences for speed level or control mode were observed for the three dimensions of the SART questionnaire: attentional demand (2.6 ± 0.2), attentional supply (4.0 ± 0.2), and understanding (5.2 ± 0.2). The rating scale was 1–7 (where 1 was “low” and 7 was “high”).

3.3.4. Motion Sickness

Mixed linear models revealed a significant difference between the control modes for central \( F_{1,47} = 5.46, df = 47.487, p \approx 0.025 \) and sopite-related \( F_{1,47} = 4.07, df = 47.430, p \approx 0.049 \) symptomology. Participants reported significantly lower amounts of central and sopite-related symptoms when the system was engaged (central: 11.3 ± 0.1; sopite-related: 15.9 ± 0.7) than when the system was disengaged (central: 11.9 ± 0.3; sopite-related: 19.9 ± 2.4). However, scores on all subscales [including gastrointestinal (11.1 ± 0.0) and peripheral (12.8 ± 0.8) subscales] were extremely low and consequently had very little practical significance.
3.3.5. Cognitive

No significant differences in RT or accuracy were revealed for either the RT or short-term-memory tasks for speed level or control mode.

3.3.6. Trust

A mixed linear model revealed no speed level main effect for positively, negatively, or neutrally framed questions. The average response to positively, negatively, and neutrally framed questions was 3.7 ± 0.1, 3.6 ± 0.13, and 5.6 ± 0.2, respectively (higher numbers indicate greater agreement with the statement). Trust scores remained fairly constant throughout the participants’ operation of the semi-autonomous system; see mean and STD values in Figure 8.

3.4. Exit Interview

On the exit interview, when asked to compare aspects of the convoy operation between the two control modes, a majority (9 of 12) of the participants felt that it was easier to maintain the designated FD with the CAST system engaged (Figure 9). Overall, participants did report a difference in usability with regard to the scanning task or stopping exercise.
About 83% (10 of 12) of the participants felt it was easier to maintain the designated FD with the system; however, about 83% also felt it was easier to stay aligned behind the LV without the system. In general, participants were indifferent (50%-60% concurrence) as to whether the semi-autonomous system engaged made it was easier to complete the scanning task, stopping exercise, and overall convoy operation. Overall, participants also expressed indifference (58% concurrence) when asked to state whether the convoy operation was more fatiguing with or without the semi-autonomous system: this result was consistent with subjective reports of fatigue that indicated no significant difference in fatigue levels between control modes.

4. DISCUSSION

This experiment investigated the ability of a semi-autonomous mobility technology to enhance performance in convoy operations. To simulate the demands of secure mobility in convoy operations, in addition to following the vehicle ahead, participants were responsible for scanning the environment and responding to the unanticipated stopping of the vehicle ahead. Participants traversed the test course with and without the semi-autonomous technology to examine differences in performance between the control modes (manual and semi-autonomous). Incorporation of semi-autonomous technologies led to enhanced performance in aspects of situational awareness (scanning task) and convoy integrity. The semi-autonomous system reduced performance on only one measure of convoy integrity: deviation from the center driving line established by the LV. Findings from this experiment indicate that incorporating semi-autonomous systems into a convoy could lead to significantly shorter convoys with increased security capabilities as compared to manually operated convoys. Subjective measures, however, revealed an important human factors issue: even though the system led to increased performance in situational awareness and in most aspects of convoy integrity, participants felt as though they performed better without the semi-autonomous technology.

4.1. System Performance

As previously stated, incorporation of the semi-autonomous system enhanced several aspects of convoy integrity, specifically maintenance of FD and reaction to the unanticipated stopping of the LV. When the semi-autonomous system was disengaged, participants tended to increase both the FD and the variability in their FD as speeds increased. The tendency to increase FD with increasing speed is a critical safety precaution taught in civilian driving as a result of human response time. The semi-autonomous system, indifferent to human limitations and human task loading, maintained a more consistent FD even at the higher speed levels. The semi-autonomous system also enhanced participants’ ability to react to unanticipated stopping of the vehicle ahead. Participants showed significant improvement in their ability to respond to unanticipated stopping (i.e., braking distance) when the semi-autonomous system was engaged. Even with specific instructions to stop the vehicle as quickly as possible, participants were unable to respond as quickly as the semi-autonomous system. Not only was the system able to respond to unanticipated stops about 80 ms faster than the participant alone, it also/consequently resulted in shorter braking distances (about 8.5 m shorter on average). The ability of the semi-autonomous system to communicate with and respond to the vehicle ahead of it indicated that the system had the potential to reduce rear-end collisions in convoy operations and safely travel at shorter following distances (i.e., shorter convoys). The ability to travel at shorter and more consistent FDs indicated that a semi-autonomous system could potentially help mitigate the “slinky effect” (variability in the length of the convoy). The semi-autonomous system also retained a very important aspect of convoy integrity: mission completion time. When the system was engaged, participants had to correct the course of their vehicle (i.e., disengage the system, realign their vehicle to the LV ahead, and then reengage the system) on average six times per driving trial; however, the overall mission completion time did not differ significantly between the control modes.

The semi-autonomous system was, however, associated with degradations in one aspect of convoy integrity: the ability to stay aligned with the LV. When the system was engaged, participants experienced, on average, an additional 1/3 m (~1 ft) in lateral offset from the footprint of the LV. Though the increase in average offset was relatively small, it may be exaggerated by the addition of more vehicles to the convoy. This could restrict the use of this semi-autonomous system in situations that required
a tight line of vehicles (e.g., narrow roads). Improvements to the autonomous mobility lateral offset capabilities and/or the incorporation of technologies with traded, adjusted autonomous, or shared control schemes (Crandall et al., 2005; Sellner, Heger, Hiatt, Simmons, & Singh, 2006) could potentially reduce this error and lead to even greater system performance. The semi-autonomous technology in this experiment used a traded control scheme in which participants traded between manual and autonomous mode on the basis of the performance of the system (e.g., course corrections). A shared control scheme (i.e., control scheme in which humans and technology can concurrently provide input to the system) may provide better performance with regard to lateral offset. An example of one such technology would be “bias” control as demonstrated by our own laboratories (ARL and TARDEC) in Somerset, Pennsylvania, in 2007 (K. McDowell, personal communication, 2008). Bias control is a scheme that allows the participants to bias the trajectory (i.e., correct the course) of the semi-autonomous system without disengaging the semi-autonomous system. Such technologies could also help to further reduce the impact of the course correction on the operator by alleviating the need to physically disengage and reengage the system.

Finally, the semi-autonomous system was associated with enhanced local area awareness of the operator. The results indicated greater numbers of targets detected, regardless of speed. The results were consistent with the notion that the semi-autonomous capability reduced the task loading associated with driving and allowed participants to devote more of their attention to scanning their environment for targets (i.e., a secondary task). Although the actual increase in performance on the local area awareness task was relatively small (∼6%), the ability to improve target identification is critical to enhancing secure mobility. Results also revealed a trend toward faster target detection at faster speeds when the semi-autonomous technology was engaged. When considering the well-established speed–accuracy trade-off in humans (Fitts, 1954), the combination of greater accuracy at faster speeds further supports the practical significance of these findings.

4.2. Perception of Performance

Even with the evident improvements in convoy integrity and local area awareness, there was still one important human factors issue associated with the incorporation of the autonomous technology: participants felt as though their performance on the convoy operation was better without the semi-autonomous technology. This finding is critical because in operational environments, humans usually have the choice as to whether to use the technology (i.e., the ability to disengage the technology), and if they feel as though they outperform the technology, they may not be as willing to use the technology in situations in which it is designed to enhance performance. So in addition to demonstrating actual improvements in performance provided by incorporating the semi-autonomous technology, it is also important to address the human’s willingness to use the technology (Parasuraman & Riley, 1997). This portion of the discussion will address some possible causes of and solutions to the discrepancy between the objective measures of performance and the subjective perceptions of performance.

Distrust in the system is a potential reason for participants perceiving that they performed better without the semi-autonomous system. Trust in the system has been identified as a key factor in the voluntary use of automation (Lee & Moray, 1992). Participant’s results on the trust survey indicate that they were neutral on trust for the semi-autonomous system. Trust results are plausible, considering the nature of the course correction—i.e., vehicle begins to follow an unintended trajectory—and the frequency with which participants had to perform a course correction (on average, once every ∼6 min). This discrepancy between objective and subjective performance may also have been a result of how participants perceived and prioritized performance. Usability results indicated that participants perceived performance differences only in FD and lateral offset: they reported that the system led to better maintenance of FD and degraded lateral offset. Based on usability results, it appears as though participants prioritized performance on the basis of the ability of the system to stay directly behind the LV (lateral offset). Lateral offset is a critical aspect of convoy and civilian driving (i.e., staying within the designated lane and/or staying on the road). It is unclear as to whether participants would have still have felt they performed better had the semi-autonomous system lacked in another area of convoy integrity (maintaining the designated/safe FD and reacting to unanticipated stopping of the LV); however, based on the results, lateral offset appears
to be a critical aspect of perceived overall performance.

Efforts put into mitigating the system’s lateral offset and the frequency with which operators have to perform a course correction could prove helpful in enhancing system performance and changing participants’ perception of the system’s overall performance. These efforts could include improving the following capabilities of the system and/or exploring alternate control schemes. Improving the system’s following capabilities (e.g., sensing system) could directly reduce the lateral offset and frequency of course corrections. An alternative control method could indirectly improve lateral offset and reduce course corrections by increasing the ease with which participants are able to correct errors in lateral offset and trajectory of the vehicle. The semi-autonomous technology in this experiment utilized a “traded” control scheme: control over the driving task was transferred between the autonomous system and humans (Crandall et al., 2005). Participants took control when the vehicle deviated considerably from the course set by the LV; after realigning the vehicle, they then transferred control back to the autonomous system. A shared form of technology may be more beneficial in reducing the magnitude of lateral offset and frequency of course corrections because it allows the human autonomy to collaborate in controlling a system (Crandall et al., 2005). Shared control allows for exploitation of desirable human qualities (e.g., maintaining a smaller lateral offset) without sacrificing the benefit of automation (maintaining a designated FD and avoiding collisions). A shared control scheme could potentially reduce the magnitude of lateral offset and the impact of course corrections by allowing the operator to calibrate the trajectory of the vehicle in real time, without having to disengage the autonomous system.

4.3. Limitations

The age range and relatively small sample size of the sample set were based on the availability of individuals with recent (i.e., within the past 2 months) experience driving in a military convoy. The average age was greater than that of the target population; however, technology acceptance generally decreases with increasing age (Morris & Venkatesh, 2000). This indicates that a younger population may have been more proficient with the semi-autonomous system; however, it is also possible that they may be more proficient without the semi-autonomous system as well (e.g., have faster reaction times to unanticipated stops). It is recommended that the same test be conducted on a small set of volunteers with ages closer to that of the target population (21–30 years) to verify these results. With regard to sample size, though significant main effects in small samples (n < 10) generally indicate a very strong effect, it is still recommended that statistical results be verified on larger sample sizes to increase generalizability.

5. CONCLUSION

This experiment demonstrates the ability of semi-autonomous mobility technologies to control a FV and allow the operator to perform better on secondary tasks (other than vehicle control). The secure mobility function used in this experiment represented a highly demanding environment that reflected the expected loading on a convoy driver: the operators concurrently maintained a specified distance between their vehicle and a LV, responded to unanticipated vehicle stops by the LV, and completed a secondary scanning task. The semi-autonomous system improved the convoy’s ability to maintain a set separation distance and increased the following vehicle’s ability to react to unanticipated stops. The system also allowed participants to generally complete a secondary scanning task with more speed and accuracy. However, incorporation of the system degraded the ability of the vehicles to stay in the same footprint.

This experiment also highlighted a critical human factors issue associated with the incorporation of autonomy in real-world applications: the operator’s perception of performance. An operator’s willingness to use a technology is often based on his or her perception of the system. So it is important to address these concerns when developing a system. In this experiment, participants felt that, overall, they outperformed the semi-autonomous system on the simulated convoy operation. It was concluded that this perception was potentially due to the number of times they had to correct the course of the semi-autonomous system and lateral offset of their vehicle from the LV when the system was engaged.

Therefore, enhancements to overall system performance in real-world applications are achieved by considering both technological and human factors solutions. Enhancements can be achieved directly by improving the dependability of the system (e.g.,
sensory systems) and indirectly by improving the operator’s ability to interact with the system—i.e., choosing an appropriate control scheme. Improving human–machine interaction allows for exploitation of desirable human qualities (e.g., maintaining a smaller lateral offset) without sacrificing the benefit of automation (maintaining a designated FD and avoiding collisions). Future research should examine the potential of shared autonomy control schemes, such as bias control, to enhance performance of semi-autonomous technologies in convoy operations.

ACKNOWLEDGMENTS

This experiment was conducted by the Combined Arms Support Command (CASCOW) and the Tank and Automotive Research, Development and Engineering Center (TARDEC) in partnership with Lockheed Martin and the Army Research Laboratory (ARL).

REFERENCES


