Evidence from a Partial Report Task for Forgetting in Dynamic Spatial Memory

Leo Gugerty, Galaxy Scientific Corporation, San Antonio, Texas

G. Sperling (1960) and others have investigated memory for briefly presented stimuli by using a partial versus whole report technique in which participants sometimes reported part of a stimulus array and sometimes reported all of it. For simple, static stimulus displays, the partial report technique showed that participants could recall most of the information in the stimulus array but that this information faded quickly when participants engaged in whole report recall. An experiment was conducted that applied the partial report method to a task involving complex displays of moving objects. In the experiment, 26 participants viewed cars in a low-fidelity driving simulator and then reported the locations of some or all of the cars in each scene. A statistically significant advantage was found for the partial report trials. This finding suggests that detailed spatial location information was forgotten from dynamic spatial memory over the 14 s that it took participants to recall whole report trials. The experiment results suggest better ways of measuring situation awareness. Partial report recall techniques may give a more accurate measure of people's momentary situation awareness than whole report techniques. Potential applications of this research include simulator-based measures of situation awareness ability that can be part of inexpensive test batteries to select people for real-time tasks (e.g., in a driver licensing battery) and to identify people who need additional training.

INTRODUCTION

Little is understood about how people represent dynamic spatial knowledge (i.e., the knowledge of rapidly changing spatial situations needed in real-time tasks such as driving). For example, drivers must keep track of the location of nearby traffic vehicles in order to maneuver successfully and avoid hazards. Drivers' knowledge of the location of nearby vehicles is an example of what has been called situation awareness. Situation awareness can be defined as the activated knowledge a person has about a dynamic scene – knowledge that is used in performing a real-time task. While driving, operators must keep track of where they are on their route and the status of their vehicle (e.g., fuel status) in addition to nearby traffic. Research suggests that errors in maintaining situation awareness are the most frequent cause of errors in real-time tasks such as driving and flying (Hartel, Smith, & Prince, 1991; Shinar, 1993). Dynamic spatial knowledge is an aspect of situation awareness that is particularly important for tasks such as driving, flying, and air traffic control.

One reason so little is understood about dynamic spatial knowledge and the broader construct of situation awareness is that these constructs are difficult to measure. Researchers have commonly used recall measures to assess peoples' dynamic spatial knowledge and situation awareness. A common measurement technique is to have participants view a changing scene, then stop the scene at unpredictable times and have the participants recall

Requests for reprints should be sent to Leo Gugerty, 685 W. Haysville Rd., Jasper, IN 47546; after December 1998, gugerty@clemson.edu. HUMAN FACTORS, Vol. 40, No. 3, September 1998, pp. 498–508.

object locations or other information about the scene. For example, Gugerty (1997) had participants recall the locations of traffic cars after three-dimensional moving driving scenes were stopped and the screen was blanked. Endsley (1995) developed the Situation Awareness Global Assessment Technique (SAGAT), in which pilots participate in a mission in a high-fidelity flight simulator. Whenever the scene is blanked, the pilots are asked a series of questions concerning information such as the location, altitude, and fuel status of their own and other aircraft.

There is a potential problem with this recall-based approach to measuring dynamic spatial knowledge and other aspects of situation awareness. In both of the previously mentioned experiments, participants recalled multiple pieces of information about a dynamic scene after the scene was removed from the simulator screen. As participants recall initial information, their knowledge about the scene may decline because of either decay or interference in memory. A number of experiments with simple, static stimuli have shown this decline, indicating that information that is recalled later in the recall sequence is recalled less accurately (Ayres, 1966; Endsley & Smith, 1996; Lawrence & LaBerge, 1956; Scheerer, 1972; Sperling, 1960). Thus this recall technique may underestimate participants' knowledge, at least for information from simple, static scenes. One purpose of the experiment presented here was to assess whether this underestimation occurs in a complex, dynamic spatial task.

In the rest of this article, I first review research showing that accuracy in recalling briefly presented scenes was affected by order of report, as well as one study using SAGAT that found that recall accuracy was not affected by order of report. Then I describe the partial report recall technique used in the current study and compare this technique with other recall-based measures of situation awareness. Finally, the results of the experiment are presented and discussed.

Order of Report

Sperling (1960) was one of the first to point out that asking participants to recall multiple pieces of information about a briefly presented stimulus may underestimate their knowledge of the stimulus information. When participants were asked to recall as many letters as they could from an entire visually presented array (whole report), they could recall only about 4 letters of a 12-letter array (33%). However, when participants were cued immediately after array offset to recall only one of the three rows in the array (partial report), they could recall 75% of the letters in whichever row was cued. Thus the partial report technique revealed that participants had a much larger capacity for the letter arrays than was suggested by the whole report technique.

Other experiments have found that when people recall multiple pieces of information from briefly presented visual stimuli, recall accuracy is related to order of report; information that is recalled later is recalled less accurately (Ayres, 1966; Endsley & Smith, 1996). For example, Lawrence and LaBerge (1956) found this effect when participants recalled the color, shape, and quantity of geometric shapes. One experiment that did not reveal an effect of order of report on recall accuracy was conducted by Endsley (1995). Using SAGAT, she had experienced fighter pilots recall multiple pieces of information about a dynamic scene (a high-fidelity simulation of a combat mission) that had been stopped. Participants spent up to 6 min recalling information after the scene was stopped. During this time, recall accuracy was not affected by the amount of time that had elapsed since the stop.

Some possible reasons for recall accuracy having been affected by order of report in the Lawrence and LaBerge (1956), Ayres (1966), and Endsley and Smith (1996) studies but not in the Endsley (1995) study could have to do with the nature of the experimental tasks. In the first three studies, participants performed abstract laboratory tasks that were novel to them and that involved static stimuli, whereas in the last study, participants performed a complex, realistic task that was familiar to them and that involved dynamic stimulus scenes.

However, an experiment by Scheerer (1972) suggests an alternative explanation of why Endsley (1995) did not find an order-of-report effect and other researchers did.

Scheerer presented participants with a row of letters and, either before the stimulus or at various times after stimulus offset, cued them to recall letters in a right-to-left or left-to-right fashion. When the direction-of-report instructions came before or immediately after the stimulus, participants could recall in either the right-to-left or left-to-right direction and were less accurate at recalling the letters recalled later in the sequence regardless of the direction of report. However, when the direction-of-report instructions came 2000 ms after stimulus offset, no order-of-report effect was found. Participants were less accurate at recalling the letters on the right side of the row regardless of the order of report.

Scheerer explained these results in terms of the order in which participants scanned a rapidly fading visual trace of the stimuli. Scanning is assumed to involve recoding the visual information into another form (e.g., auditory working memory). Items scanned later are recalled less accurately because the visual trace has faded. In the absence of instructions, people usually scan from left to right and complete the scan within 2 s. Thus, direction-of-report instructions that are delayed 2 s are too late to affect scanning. However, if the direction-of-report instructions come before or right after stimulus offset, people can alter the order of scanning.

This interpretation implies that for a short time after the offset of a visual stimulus, people have access to fairly accurate visual information about the stimulus. However, once this visual information has been recoded into another form, the detailed visual information is lost. Thus, recall techniques that begin eliciting participant reports within the first few seconds after the offset of a visual stimulus will show order-of-report effects because items reported later will be based on information in the faded visual trace. Recall techniques that take longer to begin eliciting participant reports will not show order-of-report effects. All of the experiments that showed order-of-report effects began eliciting participant reports within seconds after stimulus offset. SAGAT participants answered probe questions for up to 6 min after each stop and took up to 13 s on average to answer each question. Thus, even the first SAGAT response after a stop may have occurred when short-term visual information was no longer available.

Experimental Task

The current experiment used a partial report technique to investigate the extent to which information in dynamic spatial memory changes during the course of recall. Participants, who were all experienced drivers, viewed moving scenes from the driver's viewpoint in a PC-based driving simulator. At an unpredictable time during each scene, the scene was stopped, and participants had to recall where nearby vehicles were located when the scene was stopped. Vehicle locations were indicated by clicking with a computer mouse on a top-down view of the road that extended ahead of and behind the driver's car. On whole report trials, participants were asked to recall cars anywhere in this response area. On partial report trials, participants were cued immediately after the moving scene ended to recall only the cars on one-third of the larger response area (just behind the driver, immediately ahead, or far ahead).

Like SAGAT, the driving task used dynamic stimulus scenes and was more realistic and familiar than the abstract laboratory tasks used in the studies that found order-of-report effects. Thus, forgetting during whole report recall would not be expected in this experiment if this effect depends on using a static, abstract, unfamiliar task. However, unlike the lengthy SAGAT recall task, the task in this experiment allowed participants to completely recall all of the car locations for a scene within about 14 s after the scene stopped, even with whole report trials. Thus, if detailed location information is forgotten within this time frame, the current experiment should detect it.

METHOD

Participants

The 26 participants were hired from temporary employment agencies and ranged in age from about 18 to 30 years. All of the participants knew how to drive.

Design

In the first session, each participant was given 72 scenes. Half of these were presented as partial report trials and half as whole report trials. In the second session, which occurred the next day, the same 72 scenes were used, but the partial report trials from Session 1 were presented as whole report trials and the whole report trials were presented as partial report trials. Within a session, the partial report and whole report scenes were not presented in separate blocks of trials. The order of scenes was randomized separately for each participant.

Of the 72 whole report scenes, 4 had three cars on the road, 22 had four cars, 64 had five cars, 32 had six cars, 16 had seven cars, and 6 had eight cars. Of the 72 partial report scenes, 24 had one car in the cued response area, 31 had two cars, 13 had three cars, and 4 had four cars.

Materials

The driving simulator software was run on personal computers (IBM compatible, 486). A detailed description of the driving simulator is presented in Experiment 1 of Gugerty (1997). The animated scenarios were shown in a window that was 15.5 cm wide by 11.4 cm high and centered on the computer screen. The duration of the animated scenarios varied randomly between 18 and 35 s. Figure 1 shows a frame from a scenario.

Participants watched each animated scenario and then recalled the ending car locations. When the animated scene ended, participants were presented with a recall grid such as the one on the left side of Figure 2. The recall grid showed the road 9 car lengths behind the driver and 17 lengths ahead of the driver. The driver's car (shown in white in Figure 2) was displayed in the correct lane. Participants could place cars anywhere on the gray part of the recall grid, not only within the grid cells. Figure 2 shows a recall grid for a partial report trial in which participants were asked to recall cars just ahead of them. In other partial report trials, the area far ahead of the driver or behind the driver was gray; on whole report trials, the entire recall grid was gray. After placing cars and clicking on the "Done" button, participants saw the grid on the right showing the correct car locations and then clicked "Done" again to start the next trial.

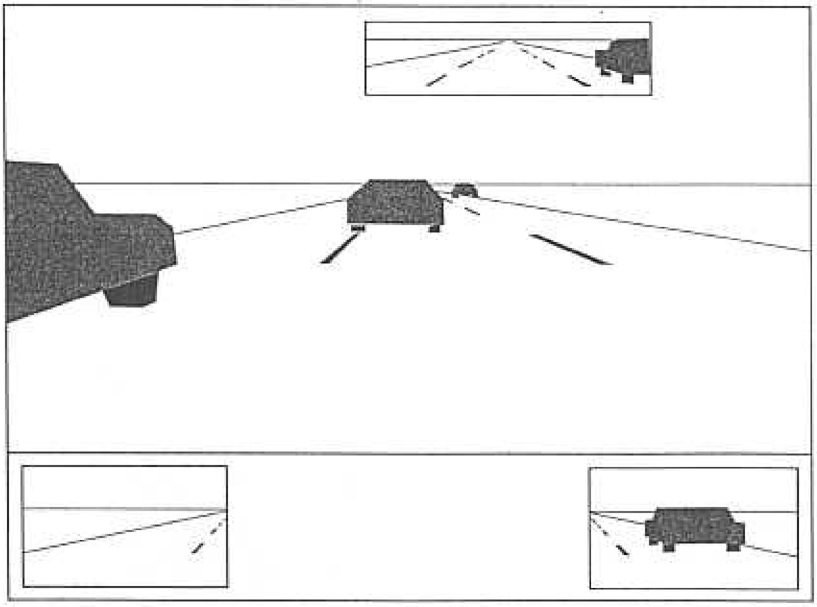

Figure 1. Three-dimensional scene from the driving simulator. The participants saw the actual scenes in color.

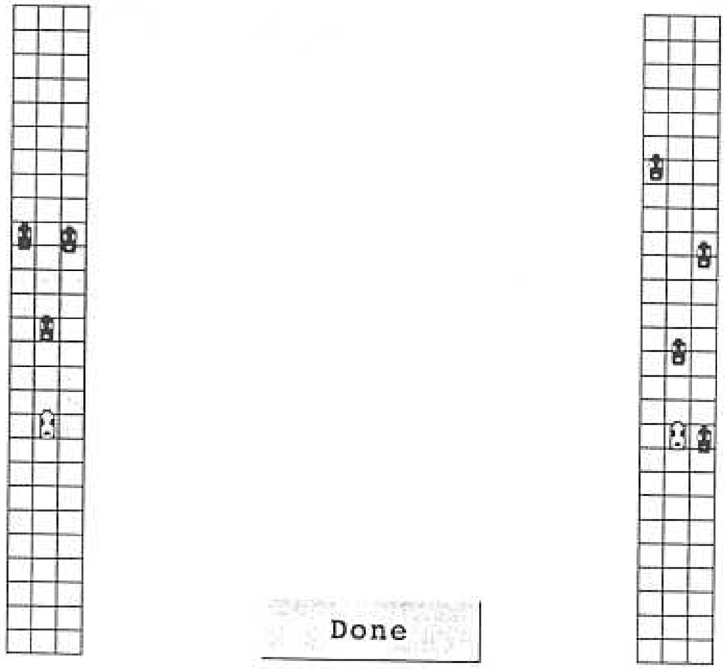

Figure 2. Recall grid (left) and feedback for a partial report trial.

Procedure

Participants were tested in groups of about 13 in a large computer laboratory. After instruction and a short practice period, each participant completed one block of 36 trials in the morning and one block of 36 after lunch, with about 1 h of unrelated cognitive tests between the block of trials. Each block took about 45 min to complete. On the following day, participants completed two blocks of 36 trials following a schedule similar to that on Day 1.

RESULTS AND DISCUSSION

Definition of Dependent Variables

Accuracy in recalling vehicles is related to both the percentage of cars recalled and the accuracy in recalling each car. In order to estimate recall accuracy, the first step was to match the cars placed by the participant with the actual cars that the participant intended to recall. For each trial, this was done using a computer algorithm that matches the participant-placed and actual cars in a way that minimizes the error distances between the participant's and the actual cars (Gugerty, 1997). In the trial in Figure 2, the participant-placed car directly ahead of the driver (on the left grid) would be matched with the actual car ahead of the driver (on the right grid), and the participant cars in the right and left lanes would be matched with the nearby actual cars in the right and left lanes, respectively. The actual car beside the driver in the right lane would be counted as an omission – a nonrecalled car. As in this example (for the car in the left lane), the matching algorithm allows a participant-placed car to be matched to an actual car outside the allowable response area. The matching algorithm also identifies intrusions – participant-placed cars that are not matched to any actual cars.

Once the participant's cars are matched to actual cars, the percentage of cars recalled and the average error distance for recalled cars can be calculated for that trial. However, participants can trade off between recalling a few cars very accurately or many cars less accurately. Therefore, to fully characterize participants' recall accuracy, a composite variable was needed that used information about the percentage of cars that participants recalled as well as information about their accuracy in recalling each car.

:ed car directly
. grid) would be
:ad of the driver
rticipant cars in
e matched with
: right and left
car beside the
e counted as an
s in this example
: matching algo-
aced car to be
le the allowable
algorithm also
ant-placed cars
ual cars.
are matched to
ars recalled and
ecalled cars can
owever, partici-
recalling a few
cars less accu-
acterize partici-
nposite variable
ation about the
ants recalled as
eir accuracy in

The composite recall error variable was defined according to the following formula,

$$E = \frac{\sum_{i=1}^{(N_c+N_o+N_i)} e_i}{N_c+N_o+N_i}, \qquad (1)$$

in which N_c, N_o, and N_i are the number of correct recalls, omissions, and intrusions, respectively. For correct recalls, e_i is the error distance between the participant-placed and the actual car. For omissions and intrusions, e_i was set at a relatively high error distance of 4.75 car lengths, which was the 95th percentile of the error distances found for correct recalls. In a previous study (Gugerty, 1997), a very similar composite variable was shown to be reliable and sensitive to experimental manipulations. The terms *hit*, *miss*, and *false alarm*, which were used in Gugerty (1997), were replaced in this report with the terms *correct recall*, *omission*, and *intrusion*, respectively, because the latter set of terms is more commonly used to describe performance on recall tasks.

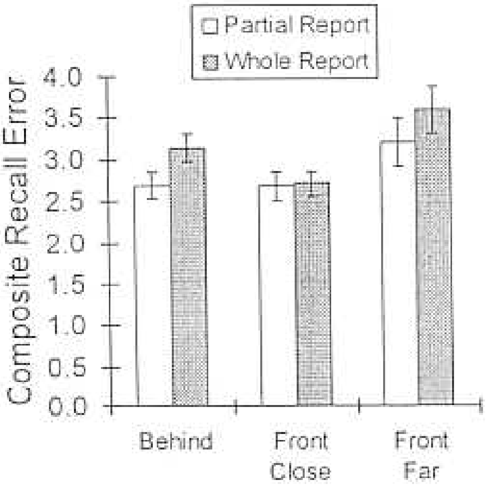

Figure 3. Composite recall error (in car lengths) for partial report and whole report trials by response area (with 95% confidence intervals).

Composite Recall Error for Partial Report versus Whole Report

For the partial report trials, composite recall error was calculated using the matching algorithm and Formula 1. The composite recall errors for each of the three response areas are shown in Figure 3, averaged over participants and the 24 partial report scenes for each response area. These data suggest that with partial report, participants were slightly better at recalling nearby cars (those behind or just in front of them) as compared with the cars farther ahead.

The main purpose of this experiment was to compare these partial report data with how well participants recalled cars in the same three response areas under whole report conditions. So, for example, I needed to calculate how well participants recalled cars behind the driver in the whole report trials. To do this, I used the 24 whole report trials that had also been shown as partial report trials with the response area behind the driver. For these whole report trials, any cars that the participants placed outside of the partial report response area – that is, any cars ahead of the driver – were dropped from the analysis. This was done because if these had been partial report trials with the response area behind the driver, participants would not have been able to place cars ahead of the driver. Then the matching algorithm and Formula 1 were used to calculate composite recall error.

Figure 3 suggests that for the area behind the driver, participants were worse at recalling cars on whole report trials than on partial report trials. Keep in mind that in making this comparison, participants' performance on the same set of scenes is being compared, with only the method of report (partial vs. whole) varying.

The same technique was used to calculate how well participants recalled cars just ahead of and far ahead of the driver in the whole report trials. Overall, as shown in Figure 3, the composite recall error was higher on whole report trials than on partial report trials for the areas behind the driver and far ahead of the driver. For the area just ahead of the driver, composite recall error was the same on whole report and partial report trials. That is, the partial report technique revealed better recall than the whole report technique for cars behind the driver and far ahead of the driver, but not for cars immediately ahead of the driver. This conclusion was supported by an analysis of variance (ANOVA), which showed a main effect of partial versus whole report, $F(1, 25) = 22.9$, $p < .001$, a main effect of response area, $F(2, 50) = 44.2$, $p < .001$, and an interaction of type of report and response area, $F(2, 50) = 3.3$, $p < .05$. Planned comparisons showed that recall accuracy was better with partial report than whole report for the response areas behind the driver, $F(1, 138) = 12.2$, $p < .001$, and far ahead of the driver, $F(1, 138) = 10.3$, $p < .002$, but not for the area close in front of the driver, $F(1, 138) = 0.10$, $p > .75$.

The overall better recall with partial report suggests that on the whole report trials, detailed location information in dynamic spatial memory is lost during recall of earlier vehicles. The lack of an advantage for partial report for the area just ahead of the driver could be attributable to participants recalling first in this area on whole report trials, which would make these whole report trials equivalent to partial report trials. The data supported this explanation. On whole report trials, the first or second car recalled was just ahead of the driver 1652 times, behind the driver 1192 times, and far ahead of the driver 503 times. The fact that drivers recalled cars right in front of them first is interesting in itself, as it suggests that drivers attend most often to this area. This conclusion was also supported by a previous experiment with the driving simulator (Gugerty, 1997).

These data suggest that information about the location of objects in a dynamic scene is lost during the course of recall. Can the data reveal anything about how quickly information is lost? The average number of cars correctly recalled per whole report scene was 5.7. Recalling all the cars on an average whole report trial took about 14 s. Over this time frame, there is evidence for loss of information in dynamic spatial memory in the driving simulator task.

Comparison of Partial versus Whole Report for Other Recall Variables

Table 1 shows how participants performed on the other two variables describing their recall, the percentage of cars recalled, and the average error distance for recalled cars. A much greater percentage of cars was recalled with the partial report (92%) than with the whole report condition (66%), $F(1, 25) = 187.7$, $p < .001$. Participants also recalled a smaller percentage of cars in the area far ahead of the driver than in the other response areas, $F(2, 50) = 28.8$, $p < .001$. In addition, the recall advantage for the partial report over the whole report condition was larger for the area far ahead of the driver than for the other areas, $F(2, 50) = 41.9$, $p < .001$. In terms of average error distance for recalled cars, participants overall were less accurate at recalling cars in the partial report (2.55 car lengths of error) than in whole report condition (2.21 car lengths of error), $F(1,25) = 21.2$, $p < .001$. However, this pattern was true only for the response areas ahead of the driver; for the area behind the driver, participants were more

TABLE 1: Recall Completeness and Accuracy for Each Response Area and Type of Report

Variable	Report Type	Response Area: Behind	Front Close	Front Far
Percentage of cars recalled	Partial report	88.8(2.1)	92.7(2.5)	93.3(2.5)
	Whole report	72.9(5.3)	81.9(4.0)	42.5(9.7)
Average error per recalled car (car lengths)	Partial report	2.36(0.15)	2.42(0.21)	2.89(0.31)
	Whole report	2.52(0.15)	2.20(0.16)	1.92(0.19)

Note: Half width of a 95% confidence interval in parentheses.

accurate at recalling cars in the partial report condition, $F(2, 50) = 23.4$, $p < .001$.

A comparison of the percentage recalled and average error distance variables shows that for the area behind the driver, participants recalled a higher percentage of cars and recalled cars more accurately with the partial report than with the whole report condition. However, for the two response areas ahead of the driver, participants recalled a higher percentage of cars with partial report but recalled them less accurately, as if they traded accuracy for completeness of recall in the partial report condition. Thus, the two variables in the table cannot be used to determine whether the partial or the whole report condition yields the best recall. For this reason, the composite recall error variable was used to compare the partial and whole report conditions.

Potential Problems Concerning the Composite Recall Error Variable

Before discussing the results further, a few potential problems concerning the preceding data analyses should be considered. The first potential problem concerns whether the composite recall error variable might be sensitive to the error distance used for omitted (forgotten) cars and intrusions. This was set at 4.75 car lengths, the 95th percentile of the actual error distances found for recalled cars. One would not want to set the error distance for omissions and intrusions lower than this, given that omissions and intrusions are errors and thus should not be assigned a lower error distance than recalled cars. So, to test whether the composite variable was sensitive to this error distance, the distance was set at the highest possible error distance for recalled cars, 8.40 car lengths (as used in Gugerty, 1997). This led to the same pattern of significant results as for the error distance of 4.75 car lengths.

A second potential problem has to do with the fact that the algorithm that matched participant-placed and actual cars allowed matches to actual cars outside of the response area. One could argue that giving participants credit for recalling a car from a noncued response area distorts the logic of partial report, as it is similar to giving credit for recalling a letter from an adjacent row in Sperling's (1960) experiment. Under close examination, this analogy does not hold up. In Sperling's experiment, the row in which a letter was located had a categorical nature in that the letters in a row did not differ in distance from the other rows. In the current experiment, however, the response area that a car is in has a continuous nature in that a car can be anywhere from the center to the edge of the area.

Consider the situation in which a participant recalled a car that was just outside of the cued response area as being just inside of it. In this situation, the participant has exhibited accurate recall. To discount the participant's accurate recall in these situations (or to count it as an error) would distort the estimate of his or her recall ability. Nevertheless, in order to see whether the conclusions of the experiment were affected by giving participants credit for recalling cars in noncued areas, I conducted an analysis in which participants were given no credit for this type of recall. This led to the same pattern of significant results as did the previous analyses.

A third potential problem concerns whether participants' performance in this experiment was at levels greater than chance. To answer this question, I used an alternative method of scoring composite recall error for whole report scenes. In all of the analyses presented previously, whole report scenes were scored in terms of how well participants recalled vehicles within particular response areas, with vehicles reported outside of that response area dropped. For comparison, each whole report scene also was scored in terms of how well participants recalled vehicles anywhere on the road (i.e., within all three response areas). In this analysis, none of the cars placed by participants were dropped. The resulting 95% confidence interval for the average composite recall error was 2.88 ± 0.15 car lengths.

In addition, to test whether participants in this experiment were recalling at levels better than chance, I estimated chance performance on whole report scenes scored in this manner. Chance performance was estimated using the method described in Gugerty (1997). The 95% confidence interval for the average composite recall error given chance performance was 3.45 ± 0.01 car lengths. Thus, participants' average composite recall error for the whole report scenes (2.88 car lengths) was much better than chance levels.

Effects of Memory Load on Recall

The previous analyses show more accurate recall under partial report conditions, except for the area just ahead of the driver, for which whole report functions like partial report. Given that it took about 14 s for participants to recall the cars in a whole report scene, the recall advantage for partial report suggests that detailed location information is lost in dynamic spatial memory over this time. An interesting question is whether this finding varies with the participants' memory load – that is, with the number of vehicles they must keep track of (traffic level). The answer to this question will give information about the capacity of dynamic spatial memory.

The memory load imposed on participants in this experiment varied from three to eight vehicles. However, scenes with three and eight vehicles were presented so infrequently that these memory loads were consolidated with the adjacent ones. Suppose that the recall advantage of partial report did vary with memory load such that, for example, the partial report advantage occurred only at higher memory loads of six or seven cars. This would suggest that people can maintain detailed spatial location data for up to five objects over the 14 or so seconds that it takes to recall a whole report scene. If, on the other hand, the recall advantage of partial report does not vary with memory load, this would suggest that even with a memory load of four or five objects, detailed spatial location information is lost within 14 s.

To answer this question, one must estimate participants' recall accuracy using the variable of percentage of cars recalled instead of the variable of composite recall error. Composite recall error was not used, given that it would create an unavoidable confound between memory load and spatial density of traffic cars. As traffic level increased, the density of traffic cars within the fixed-length response area necessarily increased. This meant that recall error actually decreased at higher traffic levels, because any car placed by a participant was more likely to be near one of the more densely placed traffic cars. The greater likelihood of close matches at high traffic levels meant that the recall measures based on error distances gave a biased picture of the effect of memory load. However, the percentage of cars recalled was less affected by this confound because the wide window for counting a car as recalled (within 8.4 car lengths) meant that for all traffic levels (even low levels) it was fairly easy for a car to be counted as recalled.

To investigate the effects of memory load, I compared the percentage of cars recalled on partial report trials with the percentage of cars recalled in the same response areas on whole report trials, just as was done for the composite recall error variable. An ANOVA was conducted with two types of report, three response areas, and four memory loads. One participant was dropped from this analysis because of missing data. Participants recalled a higher percentage of cars under the partial report condition than for the same response areas under the whole report condition, $F(1, 24) = 166.6$, $p < .001$. In addition, the percentage of cars recalled declined with memory load, $F(3, 72) = 17.4$, $p < .001$. However, the partial report advantage did not vary with

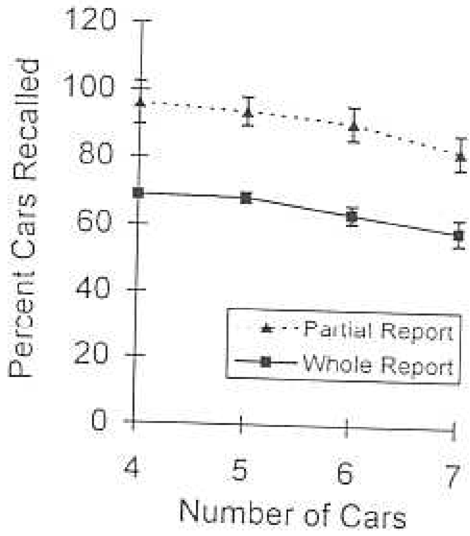

Figure 4. Percentage of cars recalled for partial report and whole report trials by memory load (with 95% confidence intervals).

memory load, as shown in Figure 4, $F(3, 72) = 0.3$, $p > .8$. Also, the decline in percentage of cars recalled with memory load was steeper for the area behind the driver than for the other response areas, $F(6, 144) = 4.4$, $p < .001$. This ANOVA also showed the same significant effects as the previous ANOVA for percentage of cars recalled – namely, a main effect of response area and an interaction of response area and type of report.

This analysis showed that participants' recall was affected by memory load. The overall decline in recall accuracy with memory load replicates findings in previous driving-simulator experiments (Gugerty, 1997). The more rapid decline in recall accuracy with memory load for the response area behind the driver suggests that at higher traffic levels, drivers allocate less attention to cars behind them. However, although memory load affected some aspects of participants' recall, it had no effect on the magnitude of the partial report recall advantage. Even at the lowest memory load used in this experiment (four cars), participants recalled more accurately with the partial report condition. Thus it seems that even at this low memory load, participants forget detailed location information over the 14 or so seconds that it takes to recall a whole report scene.

CONCLUSION

The current study showed that participants forgot information about later items in a scene but recalled earlier ones, as did the studies by Lawrence and LaBerge (1956), Sperling (1960), Ayres (1966), Scheerer (1972), and Endsley and Smith (1996). This study extends those findings because it used dynamic, realistic scenes instead of static, abstract stimuli. Also, the current study showed that participants who had considerable relevant task experience forgot information in dynamic spatial memory, whereas the other five studies all used naive participants.

The current study does not fit with the results of Endsley's (1995) SAGAT study, which showed that experienced pilots performing a high-fidelity dynamic task could recall later items as accurately as earlier ones over a 6-min period. Thus the current study showed that the lack of an order-of-report effect in the SAGAT study was not attributable to the realistic, dynamic nature of the task or the experience level of the participants.

The key difference between the current study and the SAGAT study was in the time frame for participants' responses. In the current study, participants usually completed recalling spatial location information within 14 s after the dynamic scene was stopped; in the SAGAT study, participants took up to 13 s to recall the first of 26 items. This suggests that even experienced participants forget detailed spatial location information soon after a dynamic scene is stopped.

The current study has implications for situation awareness measurement techniques, such as SAGAT, that rely on probing operators' memories. Although SAGAT may reveal considerable useful information concerning the knowledge used by real-time operators, by waiting so long to assess this knowledge it may miss some of their more fine-grained knowledge, such as detailed spatial location knowledge. The partial report technique, by limiting the time for participants' responses, is a useful method for assessing this fine-grained, short-term knowledge.

The magnitude of the information loss during whole report recall that was found in this study seems large enough to have practical significance. Ignoring the response area just in front of the driver, for which the whole report trials tended to function like partial report trials, the participants recalled cars about half a car length less accurately on whole report trials than on partial report trials. The analysis based on number of traffic cars showed that this information loss occurred even when the memory load was as low as four cars.

ACKNOWLEDGMENTS

This research was performed at the Armstrong Laboratory, Cognitive Technologies Branch, Brooks Air Force Base, Texas, with support from the Air Force Office of Scientific Research (Grant PE 61102F, Work Unit

2313T147, AFOSR Task 2313/BA [LAMP], LRIR#92AL009). The research was performed under a fellowship from the National Research Council. I thank William Tirre for helpful comments at all phases of the research, Bob Young, Doug Eddy, MaryLou Cheal, Richard Fahey, and two anonymous reviewers for comments on the manuscript, Roger Dubbs and Jamie Burns for programming the driving simulator, and Brennan Underwood and Linda Robertson-Schule for help in data collection.

REFERENCES

Ayres, J. (1966). Some artifactual causes of perceptual primacy. *Journal of Experimental Psychology, 71*, 896–901.

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

Endsley, M. R., & Smith, R. (1996). Attention distribution and decision making in tactical air combat. *Human Factors, 38*, 232–249.

Gugerty, L. (1997). Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory. *Journal of Experimental Psychology: Applied, 3*, 42–66.

Hartel, C., Smith, K., & Prince, C. (1991). *Defining aircrew coordination: Searching mishaps for meaning.* Paper presented at the Sixth International Symposium on Aviation Psychology, Columbus, OH.

Lawrence, D., & LaBerge, D. (1956). Relationship between recognition accuracy and order of reporting stimulus dimensions. *Journal of Experimental Psychology, 51*, 12–18.

Scheerer, E. (1972). Order of report and order of scanning in tachistoscopic recognition. *Canadian Journal of Psychology, 26*, 382–390.

Shinar, D. (1993). Traffic safety and individual differences in drivers' attention and information processing capacity. *Alcohol, Drugs and Driving, 9*, 219–237.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 74*(11), 1–29.

Leo Gugerty received a Ph.D. in experimental psychology from the University of Michigan in 1989. He is a research psychologist at Galaxy Scientific Corporation, San Antonio, Texas.

Date received: October 14, 1997
Date accepted: February 24, 1998