Effects of capacity demands on picture viewing

JAMES R. ANTES and ARLINDA F. KRISTJANSON
University of North Dakota, Grand Forks, North Dakota

Effects of cognitive-resource demands on picture-viewing patterns were investigated. The eye fixations of 72 subjects were recorded as the subjects viewed pictures and concurrently performed one of three listening tasks. Half of the subjects were asked to remember certain objects from the pictures and half had no-memory instructions. Concurrent auditory monitoring increased interfixation distances and the frequency of fixations on regions of high informativeness, and decreased the area of the pictures explored and the memory for objects in the scenes. It is suggested that the demands on cognitive resources influenced subjects' ability to encode and integrate fixated information and therefore prolonged the normal first phase of viewing, described by Buswell (1935).

The pattern of eye movements exhibited by observers during the free viewing of pictures has been well documented (Antes, 1974; Buswell, 1935; Yarbus, 1967). Buswell described two phases of viewing. The first involves a broad survey of the scene, characterized by short fixation pauses. This is followed by fixations of longer durations on more localized areas of the picture. Antes found that during the first phase observers tend to make relatively long saccades to areas high in information content (rated informativeness). Interfixation distances are shorter during the second phase, during which the viewer tends to fixate on less informative details.

The distribution of eye fixations during picture viewing appears to be an obvious example of selective attention, with the point of fixation operationalized as the locus of attention. According to this approach, some cognitive mechanisms can be applied to only one task or input at a time. Beginning about two decades ago, capacity models of attention (Kahneman, 1973; Moray, 1967) emerged to provide another way of thinking about attention. Attention was conceived as a limited pool (or set of pools) of resources that could be allocated to different inputs or tasks. Researchers began to think in terms of how much of the pool of cognitive resources was drawn by various inputs or phases of a task, in addition to considering the cognitive mechanisms that influenced which inputs we attend to.

Although attempts have been made to integrate the two concepts of attention (Dark, Johnston, Myles-Worsley, & Farah, 1985), no research has been done to investigate the capacity demands associated with picture viewing. Some mental effort is presumably involved in the pro-

cesses by which the viewer gathers information from the region around the fixation point, makes a decision about where to fixate next, and builds an evolving representation of the scene. The study of the capacity associated with these processes may provide a more complete understanding of the phases of picture viewing described by Buswell (1935) and Antes (1974).

The purpose of the present research, then, was to investigate the capacity demands associated with picture viewing. The strategy involved engaging the subject in different tasks that varied in complexity while pictures were simultaneously being viewed. The extent to which normal viewing patterns were influenced by the concurrent tasks was an indicator of the capacity normally required in viewing pictures.

The eye movements of observers were recorded as they viewed a series of line drawings. A third of the drawings were viewed while the subjects performed a simple listening task, a third while performing a complex listening task, and a third while no concurrent task was performed.

We were concerned that the requirement to perform a concurrent task would encourage a strategy of not examining the pictures at all. Therefore a brief memory task was introduced, which enabled us also to investigate whether or not resources drawn away from the task of viewing the pictures would influence what the subjects could remember from the pictures. After each drawing was viewed, the subjects were shown a representation of the same scene, in which objects were replaced by geometric shapes. Four of the shapes were designated by letters, and the subjects were asked to label the objects from the scene designated by the letters.

The introduction of the memory task produced a further complication, however. The study of eye movement patterns during picture viewing has largely occurred under "free viewing" instructions, without a memory task. If memory instructions influence viewing patterns, then effects of the concurrent tasks would be confounded with

The authors wish to express appreciation to Tandi Brayson, Jennifer Laabs, and Robin Michael for their assistance in data collection. Requests for reprints should be addressed to J. R. Antes, Department of Psychology, Box 8380, University of North Dakota, Grand Forks, ND 58202.

effects of the memory instructions. Indeed, instructions given prior to the presentation of the material to be viewed has been shown to influence eye movement patterns in such tasks as reading (e.g., Grabe, Antes, Thorson, & Kahn, 1987), picture viewing (e.g., Buswell, 1935; Yarbus, 1967), and examining X rays (e.g., Kundel & Nodine, 1978). Particularly relevant is a study by Friedman and Liebelt (1981), who asked observers to examine a series of line drawings in preparation for a difficult recognition memory test. They found that subjects distributed their eye fixations evenly across all the objects in the drawings, regardless of the informativeness of the objects. Therefore, the variable of instructions was introduced. Half the subjects received a memory test following each picture presentation, and half did not.

The cognitive processes that are presumably required in scanning a picture include (1) encoding the fixated information, (2) integrating information obtained during the current fixation with information from previous fixations, (3) processing information peripheral to the point of fixation, and (4) making a decision regarding where to look next. Effects on the encoding process would be indicated by deficits in memory for fixated objects as a function of the complexity of a simultaneous listening task. Evidence for disruption of the integration process by the listening tasks would be a lengthening of Buswell's (1935) first viewing phase, which is demonstrated by relatively long interfixation distances intervening between relatively short fixational pauses on highly informative regions. Effects on processing peripheral information would be revealed by changes in the "useful field of view" (Mackworth, 1965). One way to operationalize the useful field of view is to measure mean interfixation distance, with the assumption that larger interfixation distances indicate processing of information farther into the visual periphery. Influences of the listening tasks on the decision of where to look next would be revealed by the distribution of fixations across listening-task conditions.

METHOD

Subjects

The subjects were 36 male and 36 female undergraduate psychology students who participated for course credit. They all reported normal vision without glasses or wore soft contact lenses.

Stimuli

There were a total of 12 complex line drawings used in this experiment, representing various indoor and outdoor scenes, selected from a larger pool of 30 drawings. Several pilot studies were conducted to select the final scenes for the experiment. For each of the drawings from the larger pool, all objects of intermediate size (approximately 2° – 4° of visual angle) were identified. In the first pilot study, 47 subjects were shown each of the drawings and asked to make two probability judgments for each object. They were asked to rate on a scale from 0 to 100 the percentage of time they would expect the object to be present in (1) "this or similar scenes," and (2) "completely different scenes." Objects with high "similar" ratings and low "different" ratings were designated as *diagnostic* objects. Objects with relatively high ratings for both similar and different scenes were called *nondiagnostic* objects. The 12 drawings

selected each contained two objects that were readily classified as diagnostic (mean ratings of 80.1% for similar scenes and 13.2% for different scenes) and two objects classified as nondiagnostic (mean ratings of 78.6% for similar scenes and 36.5% for different scenes). Figure 1 shows one of the scenes with the diagnostic and nondiagnostic objects designated.

The 12 scenes were divided into three equivalent sets of four drawings, based upon the difficulty a second pilot group of 20 subjects had in remembering the four critical objects (diagnostic and nondiagnostic). Informativeness ratings were obtained for each of the drawings in a third pilot study involving 19 subjects. A 4×4 matrix was drawn over each picture, creating 16 equal-sized square sections. The subjects were asked to rate the informativeness of the sections by rank ordering, from 1 to 16, with the information conveyed by each section ranging from *least informative* (1) to *most informative* (16). Mean rankings were obtained for each section and used later in the analysis of eye fixation patterns.

Listening Tasks

There were three listening tasks: no task, a simple task, and a more complex task. The subjects experienced all three listening tasks, each during one third of the trials. For the simple listening task, the subjects heard a recorded list of two-digit numbers presented at a rate of one number per second and were asked to press a key when they heard a number evenly divisible by five. For the complex listening task, the subjects were asked to press a key when they heard three consecutive odd numbers. The number of times a keypress was appropriate was equated for the two tapes and was approximately five per 20 sec.

The data from the present study support the selection of these listening tasks as differing in amounts of cognitive resources required for successful performance. Kahneman (1973) argued that pupil diameter is a measure of the cognitive resources that are engaged during the performance of a task. Mean pupil diameter was determined for each subject under the three listening-task conditions across memory instructions, and a one-way listening-task (none/simple/complex) analysis of variance was computed. The

Figure 1. One of the drawings used in this research, with the two diagnostic (solid lines) and nondiagnostic (dashed lines) objects outlined.

listening task effect was significant [$F(2,142) = 99.70, p < .001$]. Subsequent Tukey HSD analysis revealed that mean pupil diameters while performing no listening task (4.04 mm), the simple listening task (4.37 mm), and the complex listening task (4.46 mm) were all significantly different from each other.

Apparatus

Eye movements were recorded with a Gulf and Western Eye View Monitor (Model 1994S). The instrument uses the relative location of the corneal reflection and pupil center to determine eye position, which is sampled at a 60-Hz rate. The 1/60-sec samples were recorded on line by a PDP-11/34 computer and converted later to eye fixations and saccades with a program modified after one described by Kliegl and Olson (1981).

Procedure

The subjects were told that their eye movements would be recorded while they were viewing a series of line drawings. The subjects who had memory instructions (half the males and females) were asked to look at the drawings carefully because they would be asked to identify some of the objects. A sample picture was then shown, followed by a sample memory slide containing the geometric shapes and letters marking objects from the drawing to be identified. The subjects who did not have memory instructions were simply asked to look at the pictures. All subjects were also told that while viewing the drawings they would be performing a listening task simultaneously. Specific instructions for each listening task were given just before performance on the tasks was required.

Each subject then saw the 12 drawings, one set of four (as defined above) while performing the simple listening task, one set while performing the complex task, and one set while performing no listening task. Across subjects, each picture set was matched with each listening task an equal number of times. Also, each listening task and each picture set appeared in each ordinal position (first, second, or third) an equal number of times during the experiment.

Each picture was shown for 20 sec. The first picture of each set was considered practice, although the subject was not told so, and the data were not recorded. The tape containing the numbers for the listening task was started about 2 sec before the drawing appeared and continued for about 1 sec after the drawing disappeared from the screen. A research assistant controlled the tape and recorded the subject's keypressing responses to the numbers on a sheet containing a transcript of the recording. After each drawing, subjects having memory instructions were given the memory test. The subject had as much time as he/she wished to write the names of the objects represented by the four labeled shapes.

All drawings were prepared as slides and projected onto a screen approximately 1.25 m from the subject. At that distance, the drawings subtended a visual angle of 16° vertically and horizontally.

RESULTS

Preliminary analyses were conducted to determine the effect of memory instructions on the eye movement measures. Since there were only minor effects involving instructions, the data were combined across instruction conditions for the following analyses.

Effects of Listening Tasks

To examine the effects of the listening tasks, the first 10 fixations, the middle 10 fixations, and the last 10 fixations each subject made on a drawing were analyzed. For each of these blocks of fixations, the mean rated informativeness of regions fixated, the mean fixation duration, and the mean interfixation distance were determined. The data are depicted in Figure 2 and were subjected to

identical listening task (none/simple/complex) $\times$ fixation block (first/middle/last) analyses of variance with repeated measures on both variables.

For the informativeness variable, there were two significant main effects and no interactions. Both listening

Figure 2. Mean informativeness of fixated regions (a), mean duration (b), and mean interfixation distance (c) as a function of fixation block for the three listening-task conditions.

task [$F(2,142) = 3.51, p < .05$] and fixation block [$F(2,142) = 7.10, p < .01$] were significant. Subsequent Tukey HSD analyses revealed that fixations during the simple listening task were directed to significantly more informative regions than were those during the no-listening-task condition. The difference between the complex listening task and no listening task was in the same direction but was not significant. There was no difference between the simple and complex listening tasks. Fixations during the first fixation block were directed to more informative regions than were those during the middle or last block, which did not differ.

In the analysis of mean duration, the main effects of both fixation block [$F(2,142) = 7.03, p < .01$] and listening task [$F(2,142) = 8.38, p < .001$] were significant. Mean durations were shorter during the first fixation block than either the middle or the last block, which did not differ. Longer mean durations occurred during the complex listening task than during the simple listening task or no listening task. The latter two conditions did not differ.

For interfixation distance, there was a significant main effect of listening task [$F(2,142) = 4.42, p < .05$] and a significant fixation block $\times$ listening task interaction [$F(4,284) = 7.88, p < .01$], which was characterized by a significant decrease in interfixation distance over the first two fixation blocks for the no-listening-task condition, a significant increase for the complex listening task, and no change for the simple listening task. When combined over fixation block, this resulted in significantly longer interfixation distances during the complex listening task than in the no-listening-task condition. Interfixation distance during the simple listening task was also greater than that during the no-listening-task condition, but the difference only approached significance.

Interfixation distance indicates how large the saccades were, on the average, but does not indicate to what extent the entire picture was viewed. It is thus possible to have many large saccades, yet to focus fixations on only a few areas of the picture. To get an indication of the extent to which the entire picture was scanned, an analysis of *regions viewed* was made. This simply represents the number of the 16 picture sections viewed at least once during the 20-sec exposure. A one-way listening-task (none/simple/complex) analysis of variance was performed on these data, resulting in a significant effect for listening task [$F(2,142) = 37.02, p < .001$]. Tukey HSD analysis revealed that all three means were significantly different from each other. When no listening task was involved, the subjects viewed 11.84 regions (74.0%); for the simple listening task, they averaged 10.68 regions (66.8%); during the complex listening task, they viewed 10.00 regions (62.5%). Thus, as greater cognitive capacity was engaged, less area of the pictures was viewed.

Memory for the Objects

To score memory performance, a key was devised by asking an independent group of subjects to label the criti-

cal objects. Figure 3 presents the number of objects correctly remembered by subjects in the present study who had memory instructions. A listening task (none/simple/complex) $\times$ diagnosticity (diagnostic/nondiagnostic) analysis of variance was computed, with repeated measures on both variables. There was a significant diagnosticity effect [$F(1,35) = 20.17, p < .001$], indicating superior memory for the diagnostic objects. There also was a significant listening-task effect [$F(2,70) = 35.20, p < .001$]. The Tukey HSD analysis showed that memory was superior in the no-listening-task condition to that in both the simple- and the complex-listening-task conditions, which did not differ.

An analysis was undertaken to determine whether the memory differences were related to the distribution of eye fixations. Loftus (1972) demonstrated a positive relationship between the number of eye fixations and recognition memory for pictures. The number of eye fixations on the diagnostic and nondiagnostic objects was determined for both instruction conditions, and a listening task (none/simple/complex) $\times$ diagnosticity (diagnostic/nondiagnostic) analysis of variance was computed. The diagnosticity effect was significant [$F(1,71) = 12.09, p < .001$], with diagnostic objects receiving an average of 1.91 fixations, and nondiagnostic objects, 1.68 fixations. The listening task $\times$ diagnosticity interaction was also significant [$F(2,142) = 3.73, p < .05$]. Listening demand had little effect on the number of fixations on diagnostic objects (1.84, 2.03, and 1.85 fixations for no listening task, the simple listening task, and the complex listening task, respectively) but did tend to reduce the number of fixations on nondiagnostic objects (1.90, 1.68, and 1.46 fixations). This interaction is consistent with the finding reported earlier that viewing becomes restricted to more informative areas with increased task demands, given that the areas encompassing diagnostic objects received higher informativeness rankings than did areas with nondiagnostic objects. Also, the fact that memory for diagnostic objects declined with listening demand but that the number of fixations to diagnostic objects remained constant suggests that the mental resources drawn during the listen-

Figure 3. Proportion of diagnostic and nondiagnostic objects correctly identified as a function of listening task.

ing tasks affected the amount of information gathered during an eye fixation.

Listening-Task Performance

Performance on the simple and complex listening tasks was assessed by examining d' values. The subjects performed significantly better on the simple listening task ($d' = 3.81$) than on the complex listening task [$d' = 3.00$; $t(71) = 5.77$, $p < .001$].

DISCUSSION

The subjects in this experiment were asked to view scenes while simultaneously performing listening tasks designed to draw cognitive resources. The cognitive tasks varied in difficulty, as indicated by d' values, and demanded mental effort, as indicated by pupil diameter. Of central interest were the effects of performing these listening tasks on the scanning patterns of subjects as they viewed the drawings.

One major effect occurred in the number of regions of the picture fixated. As listening demand increased, fewer regions of the picture were examined. The informativeness analysis indicated that the regions that were viewed during concurrent performance of the simple and complex listening tasks tended to be those ranked as highly informative. The listening tasks also had an effect on interfixation distance. When no listening task was being performed, the subjects showed the typical tendency to have, on the average, relatively large saccades early in viewing and relatively smaller saccades as viewing progressed. When the simple listening task was being performed, interfixation distance remained at the relatively longer level characteristic of early viewing throughout the viewing period. During the complex listening task, interfixation distance actually increased from the level present early in viewing.

The listening tasks also degraded the memory for both diagnostic and nondiagnostic objects in the drawings. Still, in each listening-task condition, memory for diagnostic objects was superior to that for nondiagnostic objects. When fixation location was related to memory performance, the results were consistent with the hypothesis that the demands for cognitive resources resulted in a reduced amount of information processed per eye fixation.

The finding that the pattern of eye movements changed as a function of concurrent task demands suggests that mental effort is required for picture viewing. The process that was most clearly disrupted, and thus may require the greatest resources, was the encoding of fixated information. This is indicated by decreased memory for diagnostic objects on simple- and complex-listening-task trials but no change in frequency of fixations on diagnostic objects.

The data are consistent with the contention that the process of integration of information across fixations also required resources and was disrupted. This is suggested by the pattern of fixation locations and interfixation distances

characteristic of early segments of normal viewing, which was demonstrated throughout the 20-sec exposure when the simple and complex listening tasks were also being performed.

Peripheral processing is the third picture viewing process that was identified above, and it appears not to have been disrupted. This is evidenced by the long saccades to highly informative areas occurring during the simple- and complex-listening-task conditions. It is worth emphasizing that, at the same time as longer interfixation distances occurred, the fixations were placed selectively to more limited areas of the scenes, the highly informative areas. This finding supports the distinction made by Hockey (1970) that engaging cognitive resources causes "attentional narrowing" and not necessarily "perceptual narrowing."

The fourth process identified above is that of making a decision regarding where to look next. As listening task complexity increased, fixations were concentrated on more restricted areas of the scenes, areas that were, on the average, highly informative. This, of course, does not necessarily indicate that the decision-making process was disrupted, and it may, instead, suggest that the process was working well. That is, the distribution of fixations is consistent with a decision-making heuristic to fixate the most highly informative area that is not fully processed.

Taken together, these results suggest that the cognitive resources taken from normal picture viewing by the listening tasks resulted in a retardation or prolongation of the normal viewing process. The larger interfixation distances and tendency to fixate highly informative regions characteristic of early phases of free viewing (Antes, 1974) continued throughout the entire exposure when subjects were performing the simple and complex listening tasks. There was an apparent pattern for subjects, when faced with the concurrent listening tasks, to fixate repeatedly the more informative regions. The results also suggest that each fixation provided less information during the simple and complex listening tasks than when no listening task was being performed. This occurred in spite of the fact that mean fixation duration increased with increasing task complexity. Loftus (1983) proposed that an eye fixation lasts at least long enough to encode the fixated picture features. The simple and complex listening tasks interfered with this encoding process. In terms of the phases of picture viewing that Buswell (1935) first described, perhaps the initial phase serves to provide the viewer with a basic general understanding of the picture content that serves as a guide for later viewing. When cognitive resources are diverted to another concurrent activity, that understanding is delayed, and the viewer's progression to the second phase of viewing is delayed.

One implication of the present research is that viewers in situations of reduced available cognitive resources are less able to respond quickly to the changing demands of the visual environment. However, the subjects were not asked to respond to the scenes, and the pictures were rather mundane, far from the situation presented to, say,

the operator of a motor vehicle. Also, the present subjects were instructed not to interrupt their performance of the listening tasks as they viewed the drawings, a factor also limiting the generalizability of these findings. Further complicating definitive conclusions from this study is the fact that eye movement measures that are aggregates of behavior over a period of time were used to make inferences about processes that presumably occur during each eye fixation.

REFERENCES

- ANTES, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology*, **103**, 62-70.
- BUSWELL, G. T. (1935). *How people look at pictures*. Chicago: University of Chicago Press.
- DARK, V. J., JOHNSTON, W. A., MYLES-WORSLEY, M., & FARAH, M. J. (1985). Levels of selection and capacity limits. *Journal of Experimental Psychology: General*, **114**, 472-497.
- FRIEDMAN, A., & LIEBELT, L. S. (1981). On the time course of viewing pictures with a view towards remembering. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 137-155). Hillsdale, NJ: Erlbaum.
- GRABE, M., ANTES, J. R., THORSON, I., & KAHN, H. (1987). Eye fixation patterns during informed and uninformed comprehension monitoring. *Journal of Reading Behavior*, **19**, 123-140.
- HOCKEY, G. R. J. (1970). Effect of loud noise on attentional selectivity. *Quarterly Journal of Experimental Psychology*, **22**, 28-36.
- KAHNEMAN, D. K. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- KLIEGL, R., & OLSON, R. K. (1981). Reduction and calibration of eye monitor data. *Behavior Research Methods & Instrumentation*, **13**, 107-111.
- KUNDEL, H., & NODINE, C. (1978). Studies of eye movements and visual search in radiology. In J. W. Senders, D. F. Fisher, & R. S. Monty (Eds.), *Eye movements and the higher psychological functions* (pp. 241-258). Hillsdale, NJ: Erlbaum.
- LOFTUS, G. R. (1972). Eye fixations and recognition memory for pictures. *Cognitive Psychology*, **3**, 525-551.
- LOFTUS, G. R. (1983). Eye fixations on text and scenes. In K. Rayner (Ed.), *Eye movements in reading* (pp. 359-376). New York: Academic Press.
- MACKWORTH, N. H. (1965). Visual noise causes tunnel vision. *Psychonomic Science*, **3**, 67-68.
- MORAY, N. (1967). Where is capacity limited? A survey and a model. *Acta Psychologica*, **27**, 84-92.
- YARBUS, A. L. (1967). *Eye movements and vision*. New York: Plenum.

(Manuscript received March 16, 1992;
revision accepted for publication March 22, 1993.)