

Something for Nothing: Is Visual Encoding Automatic?

Annie Lang

Department of Telecommunications
Indiana University

Robert F. Potter

Department of Telecommunication and Film
University of Alabama

Paul D. Bolls

Department of Telecommunications
Indiana University

Two experiments tested the hypothesis that visual encoding of television messages is a relatively automatic process, whereas verbal encoding is a relatively controlled process. Subjects viewed 30 messages crossed on Production Pacing (slow, medium, fast) and Arousing Content (calming, arousing). It was argued that as pacing and arousal increased, the resources required to process the messages would increase, which would interfere with the controlled process of verbal encoding but not with the automatic process of visual encoding. As expected, visual recognition was not affected by the increased resource requirements, but verbal recognition declined.

TWO DIFFERENT TYPES OF COGNITIVE PROCESSING

Past research in cognitive psychology suggests that individual acts of cognition can be viewed as falling along a continuum anchored by the terms automatic processes and controlled processes (Schneider, Dumais, & Shiffrin, 1984; Shiffrin, 1988; Shiffrin & Dumais, 1981). Researchers subscribe to varying

Requests for reprints should be sent to Annie Lang, Department of Telecommunications, Indiana University, Radio TV-Center, Bloomington, IN 47405.
E-mail: anlang@indiana.edu

criteria related to the degree of consciousness involved in automatic processes. Most agree, however, that cognitive capacity limitations serve as a useful criterion for distinguishing between automatic and controlled processes (Pashler, 1998). This literature also suggests that, although no task is purely one or the other, whether a task is viewed as mainly controlled or mainly automatic depends partly on the amount of cognitive resources required to execute it. Controlled processes require more cognitive resources to complete and are therefore more susceptible to a loss in performance as a result of overall cognitive load. Tasks that approach the automatic end of the continuum, on the other hand, can be conceptualized as requiring very few cognitive resources to execute; in other words, they can be thought of as being relatively cost free within a limited capacity framework of cognition (Kahneman, 1973; Schneider et al., 1984; Shiffrin & Dumais, 1981).

Conceptualizing tasks as falling along such an automatic–controlled continuum certainly corresponds to personal experience. An example provided by Eysenck (1993) demonstrates this point. Take a moment to remember what it was like when you were learning to drive a car. Undoubtedly, you had to concentrate quite intently on the sequential steps required to do even the simplest of tasks, such as changing lanes on a freeway. You may have even found it helpful to repeat the steps and rules silently (or aloud!) to yourself: “Keep the speed constant while you check your rearview mirrors. Turn and look over your shoulder to check your blindspot. If it's clear, switch on your turn signal, and slowly veer into the next lane. Now, turn off your turn signal.” Heaven help the passenger who tried to carry on a conversation with you while you were ticking off these sequential steps in your head! You could not afford to allocate any cognitive resources to processing their conversation because you were giving most everything you had to the controlled process of driving your vehicle. The high mental load of executing the maneuver, in fact, may have caused you to ask your friend to be quiet while you were driving.

In contrast, compare that recollection to how you change lanes now: “Long sequences of movements and decisions may be carried out with minor effort and little attention” (Shiffrin & Dumais, 1981, p. 111). In other words, the task has moved along the continuum and now lies very near the automatic pole. Talking to a friend while making a lane change on the highway is no longer mentally taxing. In fact, you are probably more than capable of changing lanes while holding a cellular phone to one ear, talking to your friend on it, and still listening for the voice of the traffic reporter on your car radio in order to determine the best route home.

With this driving example it is easy to understand why the human brain as a biological system would develop the ability to automatize often-repeated tasks. Doing so allows the cost of these tasks to be eliminated or greatly reduced. This frees up processing resources, which can be allocated to controlled tasks, and allows a greater number of tasks to be completed simultaneously (Shiffrin, 1988). By moving a task toward the automatic side of the processing continuum, the cognitive system increases its ability to “multi-task.”

CONTROLLED AND AUTOMATIC PROCESSING OF MEDIATED MESSAGES

Evidence supporting the existence of primarily automatic processes has been obtained through experiments following a dual-task paradigm (Navon & Gopher, 1979, 1980). In these experiments research participants perform two tasks, such as reading a passage and taking dictation (i.e., Spelke, Hirst, & Neisser, 1976). Researchers examine how task complexity and practice interact to affect performance on the two tasks (Pashler, 1998). Change in performance in the tasks is then interpreted as evidence of how many resources are required by and allocated to the two tasks.

Several researchers have studied the processing of television using a dual task paradigm (Basil, 1994; Graber, 1990; Grimes, 1991). In these studies, television is conceptualized as a medium that presents two simultaneous, continuous streams of information—one audio and one video. The semantic correspondence (or audio/video redundancy) between these two streams of information varies both between and within messages. Researchers have suggested that as the audio/video redundancy between the audio and video messages decreases, the task of encoding the two streams of information becomes more difficult.

Grimes (1991), for example, used secondary task reaction times (STRTs) as a measure of resource allocation, and audio and video recognition tests as measures of audio and visual encoding. He found that as audio/video redundancy decreased (i.e., as cognitive load increased), recognition for information presented in the audio track decreased substantially. On the other hand, recognition of video information and secondary task reaction times elicited by visual cues remained the same or improved. Grimes attributed these findings to research participants’ ability to encode visual information automatically, while having to rely on the controlled allocation of cognitive resources to encode the auditory information. This suggestion echoes differences found between the way auditory and visual stimuli are processed (Broadbent, 1958; Kahneman, 1973).

Generally, selective attention research suggests that auditory processing requires significant processing resources. Although theoretical stances vary on the precise point in auditory processing where the “bottleneck” or “filter” or limitation exists, all theories agree that unselected or unattended auditory information (i.e., auditory information to which controlled processing resources are not allocated) receives little or no processing (Deutsch & Deutsch, 1963; Eysenck, 1993; Treisman & Gelade, 1980). On the other hand, visual processing is generally thought to be a rapid and parallel process. The early stages of visual processing are often thought to be relatively cost free (Eysenck, 1993; Treisman, 1988).

Basil (1994), in a review of the literature on attention and television viewing, concluded that many of the findings suggest that visual information is “easier” to process than auditory information. This possibility was also supported by A. Lang (1995). In her review of the audio/video redundancy literature she identified 16 studies that compared recognition memory for messages that varied on audio/video redundancy. In those 16 studies, she found nine tests of the effect of audio/video redundancy on audio recognition and eight tests of the effect of audio/video redundancy on visual recognition. All nine of the audio recognition measures yielded significant results, showing that decreasing the redundancy of the television message decreased audio recognition for the messages. Of the eight tests on visual recognition, six showed that visual recognition was not harmed by decreasing audio/video redundancy. Lang concluded that this might provide evidence for the notion that visual encoding was a relatively automatic task, whereas audio encoding appeared to be a more controlled process. She also suggested that an interesting direction for future research would be a test of the idea that “to the extent visual information is encoded automatically, it should be encoded at virtually the same level regardless of the capacity requirements of the stimulus” (1995, p. 95). Conversely, the controlled task of encoding audio information should eventually suffer as the capacity requirements of a message increase.

VARYING THE COGNITIVE REQUIREMENTS OF TELEVISION MESSAGE PROCESSING

The limited-capacity theory of television viewing (A. Lang, 1992; A. Lang, Dhillon, & Dong, 1995; A. Lang, Newhagen, & Reeves, 1996), provides a framework that can allow us to vary the relative cognitive load being placed on a television viewer by various television messages. The limited-capacity theory of television suggests that viewers are limited-capacity information processors who allocate processing resources to the task of television viewing as a result of

both controlled and automatic processes. Viewers use controlled processes to allocate resources to television viewing in response to instructions, desires, motivations, interest, difficulty, and so on. The television calls attention to itself through the elicitation of automatic resource allocation mechanisms. Many structural and content features of television have been demonstrated to elicit orienting responses in attentive television viewers (A. Lang, 1990; A. Lang, Geiger, Strickwerda, & Sumner, 1993; Thorson & Lang, 1992) and a subsequent increase in resources allocated to the message (A. Lang, Bolls, Potter, & Kawahara, in press; A. Lang, Geiger, Strickwerda, & Sumner, 1993; A. Lang, Dhillon, & Dong, 1995).

As discussed earlier, varying the degree of audio/video redundancy is one way to vary the cognitive load imposed by a television message. Other structural and content manipulations have also been identified with increasing the resource requirements of a message. Among these are the introduction of visual and auditory structural features (Reeves et al., 1985; Reeves & Thorson, 1986; Reeves, Thorson, & Schleuder, 1986), scene changes (A. Lang, Geiger, Strickwerda, & Sumner, 1993), pacing of scene changes, (A. Lang, Bolls, et al., in press), emotional valence (A. Lang, Dhillon, & Dong, 1995), emotional arousal (Bradley, Greenwald, Petry, & Lang, 1992; A. Lang, Dhillon, & Dong, 1995), content difficulty (Thorson & Lang, 1992), and narrative structure (A. Lang, Sias, Chantrill, & Burek, 1995). This study will use two of these—arousing content and pacing of scene changes—to investigate the resource requirements of audio and video encoding at various cognitive loads.

Arousing content was chosen for both practical and theoretical reasons. First, practically, television messages vary widely in terms of their potential to arouse viewers, and viewer arousal plays an important role in many aspects of information processing. Theoretically, it is argued that arousing content will increase cognitive load in a way that impacts audio and video encoding equally. In other words, the encoding of arousal is not primarily an audio or a video task.

Research has shown that messages that have arousing content require more cognitive capacity to process than messages that are relatively calm. For example, Bradley et al. (1992) found that research participants viewing arousing slides had slower secondary-task reaction times than did those viewing calming slides. Bradley and her colleagues suggested that this was due to the increased demand for resources imposed by the arousing content. Similarly, A. Lang, Dhillon, and Dong (1995) and A. Lang et al. (in press) had research participants view television messages that were calming or arousing. They also found slower secondary-task reaction times during arousing messages compared to calming messages, suggesting that the arousing messages required more resources.

The second variable used in this study to manipulate cognitive load is the pacing of scene changes. This variable was chosen because it should increase the cognitive load disproportionately on the visual encoding process. Because the prediction being made here is that visual encoding is relatively automatic, it only makes sense to stress the visual encoding process as much as possible in an attempt to locate its resource limitation. Scene changes have been shown to increase cognitive load (Geiger & Reeves, 1993; A. Lang et al., 1993; Thorson & Lang, 1992). This is particularly true when the scene change is to a completely new or unrelated visual scene (A. Lang et al., 1993). The more scene changes in a message, the more new information is presented (visually) and, as a result, the greater the cognitive load on the visual encoding system.

This experiment was designed to test the hypothesis that the encoding of visual information from a television message is a relatively automatic process, whereas the encoding of auditory information is primarily a controlled process. In other words, this theory envisions visual encoding as relatively cost free—placing little or no strain on the overall pool of cognitive resources. If visual encoding is a relatively automatic process, then we would expect it to be relatively unaffected by either of the manipulations used in this study. Neither increasing the arousing content, nor increasing the pacing of scene changes should result in a decrease in the participants' ability to encode visual information. On the other hand, if verbal encoding is resource limited, as predicted, both increasing the arousing content and increasing the pacing of scene changes should result in a decrease in participants' ability to encode verbal information. Audio and video recognition will be used to measure the audio and visual encoding processes. Encoding is chosen as the dependent variable because it is at the encoding stage that audio and visual processing appear to differ most in their resource requirements (Eysenck, 1993). Thus, recognition measures are used as the appropriate indicator of the efficiency of the encoding process (Craik & Lockhart, 1972).

The design of this experiment will cross three levels of pacing (slow, medium, and fast) and two levels of arousing content (high and low) to create messages with varying resource requirements. Messages that are the least cognitively taxing are those that have calm content and are slow paced. Within the calm messages, this yields three levels of cognitive load: calm-slow, calm-medium, and calm-fast. Similarly, within the arousing messages there are three levels of cognitive load represented by the arousing-slow, arousing-medium, and arousing-fast messages. In addition, the presence of arousing content should also increase the cognitive load. Thus, at each level of pacing, arousing messages should require more resources than calm messages. Across levels of pacing, faster pacing should require more resources.

It may be helpful to view this relative cognitive load as having three levels of magnitude: messages that are the least taxing (calm and slow), messages that are the most taxing (arousing and fast), and messages that fall somewhere in between (have arousing content or faster pacing).

Thus, overall, the prediction is for an Arousing Content (low/high) x Pacing (slow/medium/fast) x Recognition type (audio/video) interaction, such that as arousal and pacing increase, audio recognition will decrease and visual recognition will not decrease. In addition to testing this overall interaction, planned comparisons will test for significant differences between the audio and visual recognition means at each level of pacing and arousal.

EXPERIMENT I

Methodology

This experiment is a mixed 3 (Order of Presentation) x 3 (Pacing) x 2 (Arousing Content) x 5 (Message) factorial design. Pacing, arousing content, and message were all within-subjects factors. The pacing factor had three levels: slow, medium, and fast. Arousing content had two levels: calm and arousing. The five levels of message represent the five repetitions at each level of Pacing x Arousing Content. Presentation order was the only between-subjects factor. In each experiment, the participants were randomly assigned to view one of the three presentation orders.

Independent Variables

Pacing. Pacing was operationalized as the number of scene changes in a 30-second message. Scene changes were defined as camera changes from one visual scene to a completely new visual scene. Based on previous research, the levels were operationalized as follows: slow messages had 0–1 scene changes in 30 seconds; medium messages had 4–6 scene changes in 30 seconds; and fast messages had 11 or more scene changes in 30 seconds (Hibbs, Bolls, & Lang, 1995).

Arousing Content. Arousing content had two levels—arousing and calm—and was measured using the Self-Assessment Mannequin (SAM) developed by Peter Lang and his colleagues (Bradley et al., 1992; P. J. Lang & Greenwald, 1985; P. J. Lang, Greenwald, Bradley, & Hamm, 1993). The SAM is a 9-point pictorial scale designed to measure emotional response along the three

dimensions of arousal, valence, and dominance. SAM has been shown to be a valid measure of emotional response to television messages (A. Lang, Dhillon, & Dong, 1995; Morris, 1995). As a manipulation check, participants in this experiment rated their own arousal following viewing of each message using SAM. Results show a significant arousing content main effect, $F(1,46) = 150.00$, $p < .001$, epsilon squared = .760, with arousing messages having a mean of 5.03 and calm messages a mean of 3.38.

Dependent Variable

Recognition. Recognition for message content was assessed using a forced-choice, four-alternative, 120-question, multiple-choice test. Four questions were asked about each message. Two of the four questions concerned information available only through the audio track. Four random orders of the recognition test were prepared in order to control for fatigue effects. Participants' responses on the recognition test were coded as either correct or incorrect. Results are reported as percentage correct collapsed across messages at each level of Pacing x Arousing Content for visual and auditory recognition.

Stimuli

Stimuli for this experiment were 30 coherent 30-second messages chosen from a pool of 312 messages that had been taped off air from a local 55-channel cable system. Each of the 312 messages in the pool had previously been coded for number of scene changes and arousing content. The five most arousing and five least arousing messages at each level of pacing were selected for this experiment. The messages chosen ran the gamut of genres available on cable TV. They included situation comedies, dramas, cartoons, talk shows, educational shows, documentaries, game shows, and news programming. Thirty-second segments (of the messages) that made sense as stand-alone messages were chosen as the stimulus messages. The stimulus messages were edited onto video tape in three semirandom presentation orders. The orders were constructed using blocks of six messages. Each block contained one message from each of the six levels of Pacing x Content arousal. The messages for each block were chosen randomly with the constraint that each message had to appear in the first or last block at least once, across the three orders.

Participants and Procedure

Participants ($n = 49$) were undergraduates enrolled in a communications course at a large western university who received extra credit for their participation. Informed consent was obtained from all participants. Participants were told that the purpose of the study was to see how people learn from watching television. Participants viewed all 30 stimulus messages in groups of 2–6. Participants were instructed to pay attention to the messages because later they would be asked questions about the messages. Following each message, research participants rated their emotional responses using the SAM scale. After viewing all the messages, participants were given a free recall test asking them to write down all the messages they could remember viewing. Upon completion of the free recall test, participants were given the multiple choice recognition test.

Results

The general prediction that visual encoding is a mostly automatic cognitive process was initially tested by submitting the recognition data to a Recognition Type (Audio/Video) x Arousing Content (low/high) x Pacing (slow/medium/fast) x Message (5) x Presentation Order (3) ANOVA. The predicted three-way interaction among arousing content, pacing, and recognition type was significant, $F(2,92) = 14.52, p < .001$, epsilon squared = .38. The overall interaction is shown in Figs. 1 and 2 and the means are given in Tables 1 and 2. Figure 1 and Table 1 show the Arousing Content x Pacing interaction for the visual recognition data. Figure 2 and Table 2 give the Arousing Content x Pacing interaction for the audio recognition data.

Interpretation and Planned Comparisons

Looking first at the video recognition data, the prediction is that increasing the cognitive load of the messages, through the addition of faster pacing and arousing content, will not harm visual recognition. Thus, visual recognition should never drop below the level achieved in the least demanding condition (i.e., calm-slow messages). In fact, the mean recognition score for calm-slow messages is lower than the mean for any other condition. Tukey planned comparisons show that there are no significant differences among any of the visual recognition means. Thus, the increase in cognitive load imposed in this study does not appear to significantly affect visual recognition.

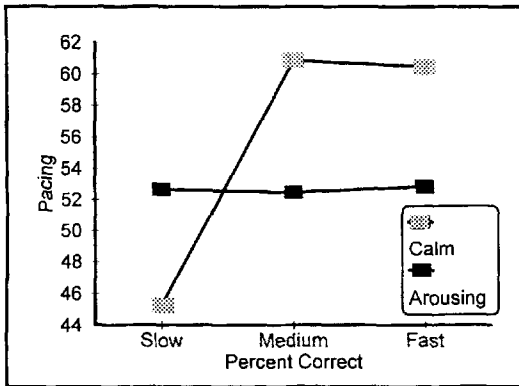


Figure 1. Visual recognition as a function of arousing content and pacing.

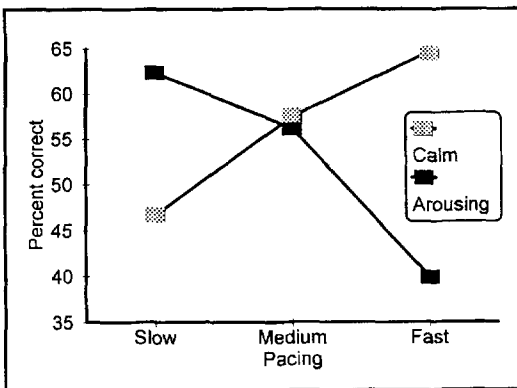


Figure 2. Audio recognition as a function of arousing content and pacing.

A very different picture appears for the audio recognition data, as illustrated in Fig. 2. For audio data, it was predicted that increasing the cognitive load of the message (by increasing pacing and arousal) would interfere with audio encoding and result in a decrease in verbal recognition. For arousing messages, this prediction is supported, with verbal recognition decreasing as pacing increases. However, for calm messages, the reverse occurs. Audio recognition for the calm-slow messages, which are the least cognitively demanding, is worse than audio recognition for arousing-slow messages and calm-medium and calm-fast messages. Thus these data partially support the hypothesis.

TABLE 1
Visual Recognition Scores for Experiments I and II

	Experiment I			
	Calm		Arousing	
	<i>M</i>	<i>n</i>	<i>M</i>	<i>n</i>
Pacing				
Slow (0–1 scene changes)	45.21 _a	36	52.67 _a	36
Medium (4–6 scene changes)	60.85 _a	36	52.48 _a	36
Fast (> 10 scene changes)	60.46 _a	36	52.85 _a	36
	Experiment II			
Slow (0–1 scene changes)	48.00 _a	30	46.00 _a	30
Medium (4–6 scene changes)	61.67 _b	30	57.00 _b	30
Fast (> 10 scene changes)	68.00 _b	30	52.33 _b	30

TABLE 2
Audio Recognition Scores for Experiments I and II

	Experiment I			
	Calm		Arousing	
	<i>M</i>	<i>n</i>	<i>M</i>	<i>n</i>
Pacing				
Slow (0–1 scene changes)	46.69 _a	49	62.36 _b	49
Medium (4–6 scene changes)	57.43 _{ab}	49	56.16 _{ab}	49
Fast (> 10 scene changes)	64.64 _b	49	39.88 _a	49
	Experiment II			
Slow (0–1 scene changes)	52.00 _a	30	68.67 _b	30
Medium (4–6 scene changes)	63.00 _b	30	54.33 _a	30
Fast (> 10 scene changes)	63.67 _b	30	46.00 _c	30

Discussion of Experiment I

The data from this experiment support the hypothesis that visual encoding is a relatively automatic process during television viewing. The two methods used to

increase cognitive load (arousing content and faster pacing) had no negative effects on visual recognition. If anything they appear to have increased visual recognition scores. Although this might at first seem somewhat troubling, it is not unexpected within the framework of the limited-capacity model of television viewing presented earlier. This theory suggests that using structural features of the medium to elicit automatic allocation of resources to encoding television messages will improve processing of the message as long as the viewer is not overloaded (A. Lang, 1995). Further, research using this model has shown that structural features that elicit orienting responses (like scene changes) also increase recognition (presumably as a result of the increase in resource allocation) for information presented 3-5 seconds following the structural feature (A. Lang, Geiger et al., 1993). Apparently, structural features increase the likelihood that information that follows them will be selected for encoding.

In this study, the calm-slow messages do not use structural features to automatically elicit additional resources. Thus, resources are primarily allocated to the calm-slow messages through controlled processes (that is, viewers attempt to comply with the instruction to pay close attention to the message.) This may mean that the increase found for medium- and fast-paced messages compared to slow messages is due to the advantage of having some information in a message "highlighted" by structural features, which then increases the likelihood of its being encoded. As long as viewers are not overloaded, this highlighting is likely to increase encoding and therefore recognition. This certainly appears to be the case for the visual recognition data, where there is no evidence of resource limitation (or overload).

Similarly, the limited-capacity model of television viewing also provides a framework for interpreting the increasing verbal recognition scores for calm messages. The theory predicts that verbal recognition will decrease at the point of cognitive overload. The data, and on reflection the theory, suggest that neither moderate arousal or moderate pacing are likely to induce cognitive overload on their own. Instead, when present singly, they simply serve to increase the allocation of resources to messages through the use of automatic resource allocation mechanisms in addition to controlled allocation. It is only when overload is achieved (which in this study appears to be when both arousing content and pacing are present in a message) that resource limitations appear and verbal recognition decreases.

Thus, the addition of just arousing content and not pacing increases audio recognition. Similarly, the addition of just pacing and not arousing content also increases audio recognition. This occurs because either variable alone results in an automatic increase in resources allocated to audio encoding but does not overload the processing system. However, the addition of both variables (pacing

and arousing content) does elicit cognitive overload, and audio recognition decreases to the lowest level.

Within this framework, the data from Experiment I do seem to support the hypothesis that verbal encoding is at least somewhat resource limited, whereas visual encoding shows no evidence of resource limitation.

EXPERIMENT II

A second experiment was conducted to replicate and extend the results of the first experiment. The same stimulus materials and recognition tests were used in the second experiment. Participants were 30 undergraduate communication majors at a large western university and were randomly assigned to presentation-order condition. The procedures for the second experiment did vary slightly from the first.

Methodology

In the first experiment, participants viewed the messages in small groups. In Experiment II, participants viewed the stimulus tapes alone seated in a reclining chair about 6 ft from a 19" television. During viewing their physiological responses to the messages were recorded using five electrodes placed on participants' forearms. (See A. Lang, Bolls, Potter, & Kawahara, in press, for results of physiological data.) Participants were instructed to stay as still as possible and to pay close attention to the television messages, because their memory for the messages would be tested later.

An attempt was made in this experiment to rule out the possibility that participants' low scores in the cognitively easy calm-slow condition were simply due to inattention or boredom. To do this, we tried to stress to the participants the importance of paying as much attention as possible to each message. Therefore, in addition to the instruction to pay close attention, participants were then shown a practice message and were given an extremely difficult recognition test on that message. The hope here was to ensure that participants were at a high level of controlled processing resource allocation.

Participants rated their emotional responses immediately following each message. After all 30 messages had been viewed, the electrodes were removed, and participants completed the recognition measures. As in Experiment I, a manipulation check was run on the arousing content manipulation. Participants rated arousing messages to be significantly more arousing than calm messages, $F(1,27) = 52.29$, $p < .001$, epsilon squared = .80. Mean arousal for calm messages was 2.81, whereas the mean for arousing messages 4.21.

Results for Experiment II

Again, the general prediction that visual encoding is a mostly automatic cognitive process and verbal encoding a relatively resource-limited process was tested by submitting the recognition data to a Recognition Type (audio/video) x Arousing Content (low/high) x Pacing (slow/medium/fast) x Message (5) x Presentation Order (3) ANOVA. The predicted three-way interaction among arousing content, pacing, and recognition type was significant, $F(2,54) = 6.42, p < .003$, epsilon squared = .16. The overall interaction is shown in Figs. 3 and 4,

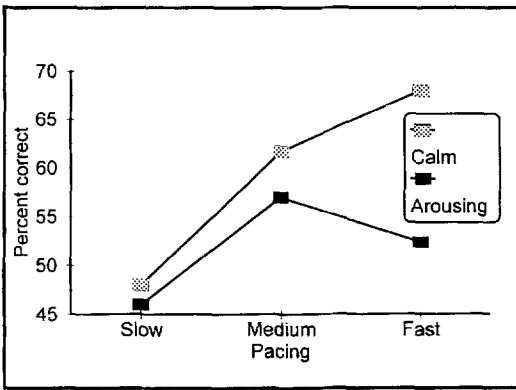


Figure 3. Visual recognition as a function of arousing content and pacing.

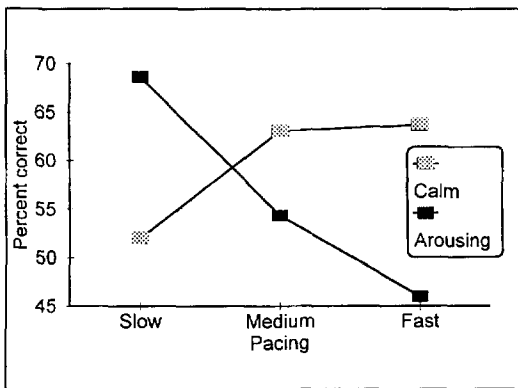


Figure 4. Audio recognition as a function of arousing content and pacing.

and the means are given in Tables 1 and 2. Figure 3 and Table 1 show the Arousing Content x Pacing interaction on the visual recognition data. Figure 4 and Table 2 show the Arousing Content x Pacing interaction on the audio recognition data. Bonferroni planned comparisons were run to compare individual means.

Interpretation and Planned Comparisons

Looking first at the visual recognition data, we remember that the overall prediction is that neither arousing content nor increasing the pacing of the message should result in a decrease in visual recognition performance. As can be seen in Fig. 3, the mean visual recognition score for calm-slow messages (the least cognitively demanding type of message) is never significantly higher than for any of the more cognitively demanding messages. Instead, as in Experiment I, increasing the pacing of the message actually increases visual recognition significantly for both calm and arousing messages. Again this pattern of results for the visual recognition data seems to support the supposition that visual encoding of television messages is relatively automatic and not severely resource limited.

Figure 4 shows the Arousing Content x Pacing interaction for the audio data. Here the prediction is that audio recognition should get worse once viewers' processing systems are overloaded. The same pattern of data appears that we saw in Experiment I. For the less cognitively demanding calm messages, the addition of pacing results in an increase in audio recognition; however, the addition of both arousing content and pacing results in a significant decline in audio recognition. Overall this pattern of data again suggests that verbal encoding is a relatively controlled or resource-limited process, but that the limitation does not appear unless both of our variables (arousing content and pacing) are present. Either one alone does not appear to overload the processing system.

Discussion of Experiment II

The results of Experiment II are very similar to those of Experiment I. The stress on paying close attention to the messages does not appear to have altered the pattern of results significantly. Similarly, the change from viewing in a group to viewing alone, which might have increased viewers' ability to concentrate on the messages, also appears to have had little effect. This similarity in results supports the hypothesis that this pattern of results is due to automatic shifts in processing resources that are not under the control of the television viewer.

CONCLUSION

These experiments draw upon the idea that mental tasks can be characterized as falling somewhere along a continuum anchored by the concepts of automatic and controlled processes. Our results support characterizing the encoding of visual information from television messages as a mostly automatic process and characterizing the encoding of audio information from television messages as at least somewhat resource limited. Consistent with the criteria that automatic processes are relatively unaffected by limitations in cognitive capacity, visual recognition never significantly declined as the cognitive load imposed by the message was increased. Even visual recognition at the highest cognitive load (fast-arousing) remained relatively unaffected by limitations in the availability of cognitive resources. On the other hand, the results show that at high levels of cognitive load, audio recognition did appear to be subject to resource limitations. Audio recognition for arousing television messages steadily declined with increased pacing, as was predicted. Interestingly, audio recognition for calm messages increased as pacing increased, in both experiments. On reflection, the limited-capacity model of television viewing provides a likely explanation for this finding. In all probability, neither arousal nor pacing alone was sufficient to overload viewers' processing capacity. Thus, the addition of just pacing, or just arousal, to messages serves to increase automatic allocation of resources to the message without overloading the processing system. As a result, audio recognition improves. However, when both arousal and pacing are present, overload occurs, and verbal recognition decreases significantly.

These results have implications that are both practical and theoretical. To begin with, our findings should be of interest to producers of television messages. Past research has shown that the arousal level, valence, and difficulty of a message all contribute to the overall amount of cognitive resources required to process it. It is hoped that wise producers have already begun factoring these findings into their preproduction decisions. However, our data suggest that another important question should be asked when planning a program concept: Will this message deliver most of its important information via the video or the audio channel? If you are counting on the audio channel to deliver much of your information, you should be especially aware of the overall cognitive load your message requires. The fact that audio encoding is a controlled process means that when viewers approach cognitive overload, the amount of resources they have available to apply to processing the auditory information decreases, and as a result their memory for what was said in the message may decline precipitously. On the other hand, because visual encoding appears to be primarily an automatic

process, information delivered via the visual channel will be much more likely to be encoded regardless of the overall resource requirements of the message. Therefore, producers may want to rely more on visual information to deliver their themes if they know that the overall cognitive load of their message will be high (i.e., highly arousing and fast paced).

From a theoretical perspective, the results of the current study provide additional evidence to support previous work on auditory and visual recognition of television messages when those messages become more difficult (Basil, 1994; Grimes, 1991; A. Lang, 1995). In addition, this study has implications for future experimental research on the processing of television messages. If, as has been suggested, visual encoding is primarily automatic, then visual manipulations used to increase the cognitive load of the message should not be expected to interfere with visual encoding. Rather the impact of these sorts of manipulations will be seen disproportionately at audio encoding and perhaps on other processes (such as storage or retrieval).

REFERENCES

- Basil, M. (1994). Multiple resource theory I: Application to television viewing. *Communication Research, 21*, 177–207.
- Bradley, M. M., Greenwald, M. K., Petry, M. C., & Lang, P. J. (1992). Remembering pictures: Pleasure and arousal in memory. *Journal of Experimental Psychology, 18*, 379–390.
- Broadbent, D. E. (1958). *Perception and communication*. New York: Pergamon.
- Craik, F., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671–684.
- Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review, 70*, 80–90.
- Eysenck, M. W. (1993). *Principles of cognitive psychology*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Geiger, S., & Reeves, B. (1993). The effects of scene changes and semantic relatedness on attention to television. *Communication Research, 20*, 155–175.
- Graber, D. A. (1990). Seeing is remembering: How visuals contribute to learning from television news. *Journal of Communication, 40*(3), 134–155.
- Grimes, T. (1991). Mild auditory–visual dissonance in television news may exceed viewer attentional capacity. *Human Communication Research, 18*, 268–298.

- Hibbs, H., Bolls, P., & Lang, A. (1995, May). *The medium is the memory: Using structural features to predict memory for random television messages*. Paper presented to the Information Systems Division of the International Communication Association, Albuquerque, NM.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Lang, A. (1990). Involuntary attention and physiological arousal evoked by structural features and mild emotion in TV commercials. *Communication Research, 17*, 275–299.
- Lang, A. (1992, May). *A limited capacity theory of television viewing*. Paper presented to the Information Systems Division of the International Communication Association, Miami, FL.
- Lang, A. (1995). Defining audio/video redundancy from a limited capacity information processing perspective. *Communication Research, 22*, 86–115.
- Lang, A., Bolls, P., Potter, R., & Kawahara, K. (in press). The effects of production pacing and arousing content on the information processing of television messages. *Journal of Broadcasting and Electronic Media*.
- Lang, A., Dhillon, K., & Dong, Q. (1995). The effects of emotional arousal and valence on television viewers' capacity and memory. *Journal of Broadcasting & Electronic Media, 39*, 313–327.
- Lang, A., Geiger, S., Strickwerda, M., & Sumner, J. (1993). The effects of related and unrelated cuts on viewers' memory for television: A limited capacity theory of television viewing. *Communication Research, 20*, 4–29.
- Lang, A., Newhagen, J., & Reeves, B. (1996). Negative video as structure: Emotion, attention, capacity, and memory. *Journal of Broadcasting and Electronic Media, 40*, 460–477.
- Lang, A., & Potter, R. F. (1996). Arousing Messages: Reaction time, capacity, encoding. *Communication Reports*. Manuscript under review.
- Lang, A., Sias, P., Chantrill, P., & Burek, J. A. (1992). Tell me a story: Narrative structure and memory for television messages. *Communication Reports, 8*(2), 1–9.
- Lang, P. J., & Greenwald, M. K. (1985). *The international affective picture system slides and technical report*. Gainesville: Center for Research in Psychophysiology.
- Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Evaluative, facial, visceral, and behavioral responses. *Psychophysiology, 16*, 495–512.
- Morris, J. D. (1995). Observations: SAM: The self-assessment manikin an efficient cross-cultural measurement of emotional response. *Journal of Advertising Research, 35* (Nov-Dec), 63–68.

- Navon, D., & Gopher, D. (1979). On the economy of the human processing system. *Psychological Review*, *86*, 214–255.
- Navon, D., & Gopher, D. (1980). Interpretations of task difficulty. In R. Nickerson (Ed.), *Attention and performance* (Vol. VIII, pp. 297–315). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pashler, H. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.
- Reeves, B., & Thorson, E. (1986). Watching television: Experiments on the viewing process. *Communication Research*, *13*, 343–361.
- Reeves, B., Thorson, E., Rothschild, M., McDonald, D., Hirsch, J., & Goldstein, R. (1985). Attention to television: Intrastimulus effects of movement and scene changes on alpha variation over time. *International Journal of Neuroscience*, *25*, 241–255.
- Reeves, B., Thorson, E., & Schleuder, J. (1986). Attention to television: Psychological theories and chronometric measures. In J. Bryant & D. Zillmann (Eds.), *Perspectives on media effects* (pp. 251–279). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Schneider, W., Dumais, S. T., & Shiffrin, R. M. (1984). Automatic and control processing and attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 1–25). Orlando: Academic.
- Shiffrin, R. M. (1988). Attention. In R. A. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology, Volume 2: Learning and cognition* (pp. 739–811). New York: Wiley.
- Shiffrin, R. M., & Dumais, S. T. (1981). The development of automatism. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 111–114). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Thorson, E., & Lang, A. (1992). The effects of television videographics and lecture familiarity on adult cardiac orienting responses and memory. *Communication Research*, *19*, 346–369.
- Treisman, A. (1988). Features and objects. *Quarterly Journal of Experimental Psychology*, *40*, 201–237.
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136.