Brain-Imaging Detection of Visual Scene Encoding in Long-term Memory for TV Commercials

Brain imaging, by steady-state probe topography, was used to investigate whether a distinct pattern of brain electrical activity in the left or right frontal hemispheres could identify which frames from new TV commercials would be recognized by consumers in an unannounced recognition test one week later. First, confirming previous research, video scenes held on-screen for 1.5 seconds or longer were better recognized. Second, after controlling for scene duration, it was found that video scenes that elicited the fastest brain activity in the left frontal hemisphere were also better recognized. This new finding suggests that the transfer of visual information from short-term memory to long-term memory takes place in the left hemisphere, not the right hemisphere as previously believed. Selection of visual content that produces a fast electrical response in the left-brain hemisphere should help to create highly memorable TV commercials.

CONSUMER CHOICES usually occur some time after exposure to advertisements; this is the case for mass media advertisements and also for direct-response advertisements, such as web advertisements, if the purchase response is not made instantaneously. It is possible that delayed choice may result from implicit brand awareness and brand attitude enhancement, without recall of the episodic advertising event or events that produced the enhancements (Rossiter and Percy, 1997), though this is not the view of the majority of advertisers. Most advertisers believe that explicit memory retention of traces of the advertisement is instrumental in influencing subsequent brand choice, and they therefore strive to create memorable advertisements. Visual recognition tests of key frames from TV campaigns, or illustrations from print campaigns, are an important measure in advertising tracking studies (Sutherland, 1993; Sutherland and Sylvester, 2000).

Consumers' recognition of visual excerpts from advertisements is regarded as proof of media plan reach and as minimally sufficient evidence that the advertisements "worked" (Aaker and Bruzzone, 1981; Rossiter and Percy, 1997). We know, from Starch studies, quite a lot about what types of visuals are recognizable from print advertisements (Rossiter and Percy, 1997). However, little is known about what types of visuals are recognizable from TV commercials. What has been missing is the ability to predict which video scenes from a commercial will "cut through" and be memorable. Presumably, these scenes should identify the brand and depict the essential message content (Rossiter and Percy, 1997).

A recent advancement in brain-imaging technology, called steady-state probe topography (SSPT), offers the fast temporal resolution needed to record cortical activity during exposure to a dynamic stimulus sequence, as presented by television programs and television commercials (Silbersstein et al., 1990; Silberstein, 1995; Silberstein et al., 2000). SSPT, which measures steady-state visually evoked potential (SSVEP), a brain electrical response elicited by visual stimuli, is essentially an innovative version of the well-known electroencephalographic (EEG) technology. An earlier study of EEG and memory for TV commercials (Rothschild and Hyun, 1990) is compared with ours in the Discussion. Other brain-imaging technologies, notably Positron Emission Tomography (PET) and functional Magnetic Resonance Imaging (fMRI), measure blood flow in the brain and do not provide fast enough recording and hence the temporal...
resolution necessary to adequately examine responses to other than static stimulus presentations.

With SSPT, recordings at the rate of 13 per second can be made from up to 64 sites on the scalp simultaneously. Respondents can participate individually or in groups of 10 to 12 at a time. Figure 1 illustrates what SSVEP looks like as brain images, with a typical image of a fast response (left panel) and a typical slow response (right panel) shown for a respondent group. Validated originally in neuropsychological studies of attention and memory, SSPT would appear to have great potential for the diagnosis and prediction of TV commercials’ effectiveness.

The present study was designed to examine one particular aspect of the spectrum of effects that TV advertising may produce—that of visual recognition of video scenes from commercials. The practical importance of this endeavor is clear. Should a distinct location and pattern of cortical activity be discriminable for scenes that are recognized subsequently versus those that are not, it will be possible to predict what content in a TV commercial will be visually memorable. Based on learning from this research, it should be possible to design key visuals to give a commercial very strong memorability.

The present study was designed to examine one particular aspect of the spectrum of effects that TV advertising may produce—that of visual recognition of video scenes from commercials. The practical importance of this endeavor is clear. Should a distinct location and pattern of cortical activity be discriminable for scenes that are recognized subsequently versus those that are not, it will be possible to predict what content in a TV commercial will be visually memorable. Based on learning from this research, it should be possible to design key visuals to give a commercial very strong memorability.

The theoretical importance of the study lies in the opportunity to observe what happens cortically when visual content is encoded and transferred from short-term memory (STM) into long-term memory (LTM). It is widely believed that visual, pictorial stimuli are encoded mainly in the right cortical hemisphere and verbal stimuli in the left (e.g., Kelley et al., 1998). However, the hemispheric encoding/retrieval asymmetry model, HERA, proposed by Tulving et al., (1994) hypothesizes that the left hemisphere plays a dominant role for encoding episodic memory (specific event) traces from STM into LTM, whereas the right hemisphere plays a dominant role in the retrieval of those traces from LTM in acts of recognition or recall. This study provides evidence on the encoding stage of the HERA model. Left-hemisphere encoding superiority has previously been demonstrated in studies that have used verbal stimuli Julving et al., 1994) and also for intentional learning of one type of visual stimulus, human faces (Nyberg, Cubeza, and Tulving, 1996). Ours is the first study, to our knowledge, to investigate the location of encoding during incidental learning of visual content in dynamic scenes, as in TV commercials.

THE EXPERIMENT

In overview, a single-group, within-subject, continuous recording experiment was conducted in which female shoppers watched TV commercials naturalistically placed in a TV program. They were then tested for visual recognition of static frames from the commercials after a oneweek delay. Group results are presented, as would be the case in an advertising pretest and tracking study.

Participants.

A market research company using local area random-digit dialing recruited the consumers participating in the study. Selected for the study were 35 women, aged 25 to 45 years, who were the primary shoppers for their households. The women were additionally screened to be right-handed, which is the basic indication of left-hemisphere dominance for language processing that characterizes a majority of the population. These constraints were observed as a homogeneous group in terms of gender and laterality is important for brain recording.

Apparatus.

The technical details of the SSPT apparatus and SSVEP recording are given in the Appendix. In essence, brain electrical activity from both frontal hemispheres is recorded for each consumer throughout her exposure to the TV program and commercials on a normal TV monitor (see Procedure). Computer software converts the brain activity to a continuous reaction measure that can identify visual content that elicits either fast or slow responses. The apparatus for the recognition test is a normal TV monitor on which still frames from the commercials are shown.
To provide a valid test of recognition, another set of 40 frames from other previously unseen U.S. TV commercials, also with brand frames removed, were selected to served as distractor or "true negative" stimuli, and these were randomly interspersed with the 40 test frames for which "true positive" recognition was sought. In the test, the 80 frames were presented, subject-paced, for up to 5 seconds each, with a 1-second interval between.

For the recognition test, consumers indicated via a two-button response (Yes, No) whether the frame was from a TV commercial shown to them in the TV program the week before.

RESULTS

Of interest, first, is the group's recognition performance for frames not previously seen (true negatives). Presumably, consumers look at a test frame while searching visual memory to try to find a match; if they find an apparent match, they report "Yes," and if they don't find one they report "No." Given that more than half of the scenes in the commercials were onscreen for durations of less than 2 seconds, and that 2 seconds is optimal for recognizing scenes (Potter and Levy, 1969; see also Rossiter and Percy, 1983), we should not expect that the maximum performance for scenes actually seen (true positives) would be 100 percent. Confident reporting of what you haven't seen may therefore provide an estimate of the maximum possible performance on this task. Results for the novel scenes (true negatives) are shown in the second column of Table 1. Recognition performance on the presumably easier true negatives task was 80 percent, suggesting that, for the consumer group as a whole, and for new, seen-once TV commercials, this may be the practical upper limit for correct recognition.
Table 1
Overall Recognition Performance for Visual Scenes from TV Commercials Seen vs. Not Seen

<table>
<thead>
<tr>
<th>Frames Seen Previously</th>
<th>Frames Not Seen Previously</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognized (percent)</td>
<td>52.8</td>
</tr>
<tr>
<td>(True positives)</td>
<td>(False positives)</td>
</tr>
<tr>
<td>Not recognized (percent)</td>
<td>47.2</td>
</tr>
<tr>
<td>(False negatives)</td>
<td>(True negatives)</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
</tr>
<tr>
<td>Number of subjects</td>
<td>35</td>
</tr>
<tr>
<td>Number of frames</td>
<td>40</td>
</tr>
</tbody>
</table>

The first column of Table 1 shows that the overall recognition of the test frames was approximately 53 percent, which is not significantly different from the guessing or chance level of 50 percent. However, this result is for all frames, without regard to their SSVEP responses.

It was expected, without regard to brain responses, that the exposure duration of the visual scene from which the frame was taken, that is, the length of time for which essentially the same visual content was on-screen in the commercial, would positively affect subsequent recognition of the frame. To examine this hypothesis, the exposure duration for each of the 40 test frames was measured electronically to the nearest hundredth of a second and then compared with recognition performance.

The relationship between exposure duration and recognition was significant (F = 3.95, df = 1.19, p < .05, 1-tailed) and there was a significant, positive, moderate-sized correlation between the two measures (r = .41, df = 38, p <.01, 1-tailed). Table 2 shows the percentage correct recognition for the 40 test frames dichotomized at the median exposure duration, which was 1.42 seconds, that is, just under 1.5 seconds when rounded.

Shorter-duration scenes (exposed for less than 1.5 seconds) produced below-chance recognition of 41 percent on average (one-sample t = -1.70, df = 19, p = .05, 1-tailed). Longer-duration scenes (exposed 1.5 seconds or longer) produced above-chance recognition of 63 percent on average (one-sample t = 2.19, df = 19, p < .05, 1-tailed).

Table 2
Recognition Performance for Visual Scenes: Shorter Exposure vs. Longer Exposure Duration

<table>
<thead>
<tr>
<th></th>
<th>Shorter Exposure (&lt;1.5 secs)</th>
<th>Longer Exposure (≥1.5 secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognized (percent)</td>
<td>41.1</td>
<td>62.8</td>
</tr>
<tr>
<td>Number of subjects</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>Number of frames</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

Clearly, the effect of scene duration on visual long-term memory recognition is quite strong. Again, the duration effect is independent of any brain activity that those scenes might have produced and is an important but not a new finding.

Results in the first column of Table 3 allow a test of the main hypothesis, for the target task, that visual scenes associated with fast SSVEP responses would produce high recognition whereas those associated with slow SSVEP responses would produce low recognition. This hypothesis was confirmed, with fast-response scenes correctly recognized in 59 percent of cases, versus 45 percent for slow-response scenes (paired-sample t = 4.9, df = 38, p < .0001, 1-tailed). Although significant, the absolute difference of 14 percent discrimination is not large.

The target task results were then analyzed further in terms of SSVEP responses measured in the left hemisphere and the right hemisphere separately. These results are shown in the second and third columns of Table 3. The numbers of fast-response and slow-response frames differ slightly for each hemisphere, as they were originally selected across both hemispheres. The right-hemisphere readings produced no difference in recognition performance, at about 54 percent each for fast-response and slow-response frames (paired-sample t = 0.17, df = 19, p > .10, 2-tailed), which did not differ from chance at 50 percent. However, the left-hemisphere readings produced a large difference, with 66 percent correct recognition of fast-response frames compared with only 38 percent correct recognition of slow-response frames (paired-sample t = 7.1, df = 17, p < .0001, 2-tailed).

Fast SSVEP response frame recognition was highly reliably above chance and slow SSVEP response frame recognition highly reliably below chance (both p < .05,
1-tailed). The absolute difference of 28 percent discrimination in the left hemisphere is double that obtained from the combined hemisphere readings. Also, if our estimation is correct that the ceiling of recognition for this task is 80 percent, then the 66 percent left-hemisphere result is a strong finding, representing 82 percent of possible recognition.

It is important to note that this left-hemisphere response-speed effect is independent of the earlier effect of scene duration. After duration was partialled out, the correlation between SSVEP response speed and recognition remained positive and significant for the left-hemisphere recordings (partial r = .48, df = 16, p < .05, 1-tailed) but not for the right-hemisphere recordings (partial r = -.17, df = 18, ns).

These results strongly suggest that, at least for women subjects with normal laterality, the left hemisphere is where encoding of dynamic visual scenes into longterm memory primarily takes place.

<table>
<thead>
<tr>
<th>Correct Recognition (percent) of</th>
<th>Both</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fast SSVEP response frames</td>
<td>58.7</td>
<td>66.0</td>
<td>53.8</td>
</tr>
<tr>
<td></td>
<td>(20)</td>
<td>(8)</td>
<td>(12)</td>
</tr>
<tr>
<td>Slow SSVEP response frames</td>
<td>45.3</td>
<td>37.9</td>
<td>54.2</td>
</tr>
<tr>
<td></td>
<td>(20)</td>
<td>(11)</td>
<td>(9)</td>
</tr>
<tr>
<td>Number of subjects</td>
<td>35</td>
<td>35</td>
<td>35</td>
</tr>
</tbody>
</table>

Note: Number of frames shown in parentheses

### DISCUSSION

The results of this experiment are theoretically important because they constitute the first investigation of the brain location of short-term to long-term memory transfer of content from dynamic, visual stimuli—in this demonstration, from TV commercials. This was made possible by the high temporal resolution of brainimaging recordings permitted by steadystate probe topography.

The finding that correct recognition was due to a faster brain electrical activity response only in the left hemisphere, and not the right, supports the theoretical prediction of Tulving et al.'s (1994) HERA model that STM to LTM encoding for retrieval of pictorial content, and not only verbal content, is a left-hemisphere activity. The evidence for left-hemisphere encoding has been based on verbal encoding tasks, with words as stimuli, and on the particular visual encoding task of face recognition (Nyberg et al., 1996). Previously it was assumed that pictorial encoding takes place in the right hemisphere, because that is the dominant hemisphere for retrieval of visual information (recognition of drawn objects and faces).

The present study is the first to confirm that pictorial stimuli in general are also encoded predominantly by the left hemisphere. If the left hemisphere is indeed the location for pictorial encoding, then this raises the issue of whether covert verbal labeling of pictures (Paivio, 1971; Kunen, Green, and Waterman, 1979; Pezdek and Evans, 1979) facilitates subsequent recognition of them, because the left hemisphere is also the dominant hemisphere for semantic-memory retrieval (memory for meaning) from LTM.

As discussed below, there was some suggestive evidence in the present study that the more recognizable left-hemisphere fast-response frames were more concrete and thus perhaps easier to verbally describe. Of course, full support for the asymmetric retrieval part of the model would require SSVEP recording during the recognition task, but we can regard as proven that visual LTM retrieval is right-hemisphere based (see Tulving et al., 1994; Nyberg et al., 1996).

The participants in the experiment were women, and some evidence from two PET studies with "emotional" stimuli suggests that brain laterality for men may be opposite to that of women (Cahill et al., 1996 with male subjects, and Taylor et al., 1998 with female subjects). For theoretical completeness, it would be useful to repeat the present study with left-handed women, and with men, both right-handed and lefthanded. However, these extensions would not change the findings for the important group of right-handed women shoppers used in the present study.

An earlier study by Rothschild and Hyun (1990) also investigated the relationship between brain electrical activity and recognition of visual (and copy) excerpts from commercials. Theirs was a commendable study considering that it was conducted more than 10 years ago, but ours must be regarded as more definitive. They tested nine TV commercials with right-handed women consumers, much like our study. However, they recorded brain electrical activity (EEG, with a measure of Alpha-wave blocking) from the occipital (visual cortex) region of the brain rather than the frontal memory sites that we measured, so they picked up visual "attention," or visual brain arousal, rather than visual "encoding," or memory transfer.
We are confident that we have identified the brain site where the transfer to visual memory takes place, with accordingly more powerful prediction of which scenes will be recognized.

Our own studies of different sites have shown only a weak relationship between visual cortex activity and recognition (Silberstein et al., 2000), and this is basically what their study found. Whereas attention is evidently necessary for subsequent recognition of scenes from a TV ad, it certainly is not sufficient, otherwise consumers would remember entire commercials after having paid attention to them once, which advertisers know does not happen. Moreover, their study employed a half-hour delayed recognition test whereas ours used a seven-day delay that better corresponds with the usual exposure-topurchase interval. We are confident that we have identified the brain site where the transfer to visual memory takes place, with accordingly more powerful prediction of which scenes will be recognized.

The results from the present study suggest two strategies for advertisers to make particular "key frames" in a TV commercial more memorable. One strategy is to maintain the same scene on the screen for, safely speaking, at least 2 seconds; this does not necessarily mean using a still or static shot, although a basically static shot is appropriate for a product pack or company logo (see Rossiter and Percy, 1997, and also Baker, 1999), but rather that the video frames over the 2-second period should focus on a single visual idea, such as the product in use, or an influential presenter.

Two seconds, in a thirty-seconds or shorter commercial, is a long time from the perspective of agency creative people, who tend to favor fast cuts. Fast-cut commercials, with an average shot-length of under 1.5 seconds, are overall less memorable (MacLachan and Logan, 1993). Key visuals should certainly be held for longer.

A second strategy, of course, would be to pre-test desired key visuals, perhaps in storyboard form, using the present methodology, SSPT, to ensure that they elicit a left-hemisphere, fast SSVEP response. This is recommended because it is better than relying on scene duration alone for memorability. A left-hemisphere fast response confidently indicates that the scene is going into memory.

Related to this, but obviously requiring more research, would be to learn from SSPT findings "what works" in terms of the types of visual content that spike a left-frontal hemisphere reaction. Some indications may be made at this stage, however. Contrasting the left-hemisphere fast-response frames with the others indicated that close-ups of people's faces (known to be attention-getting in print advertising research; Kroeker-Riel, 1993) are a memorable content type. On the other hand, abstract, hard-to-label, visual scenes seemed to have low memorability (for a previous review of concrete versus abstract stimuli, see Rossiter and Percy, 1983, and see the specific studies of picture recognition by Koen, 1969, and Nelson, 1971). Easy-to-label scenes could mean that the words in the TV commercial's audio may play a role in prompting the consumer to encode a visual scene; although it may be noted that the presence of on-screen words (an on-screen "super") did not appear to be a factor one way or the other in the present experiment. Also, seemingly counting against the "words help" possibility is the finding that higher-acceptance, "emotional" TV commercials seem to have their visual frames better recognized (Ambler and Burne, 1999). Obviously, we still have much to learn about what makes a visual scene memorable in TV advertising, even though we can now predict a scene's memorability.

CONCLUDING REMARKS

The brain-imaging technique of steadystate probe topography (SSPT) is the most recent of the psychophysiological techniques to become commercially available for pre-testing advertisements. SSPT brain imaging is especially promising for testing TV commercials because of fast recording during presentation. SSPT, as noted in the introduction, can record from multiple sites in the brain. Specialized sites measure consumer attention to the video and audio content throughout the commercial, as well as which parts of the content are going into visual memory, which was the focus of the present study, and which parts are going into verbal memory.

Most interestingly, we have found that attention or so-called "engagement" during the commercial does not necessarily mean that the material is being remembered, and, in fact, attention often declines while the consumer is transferring images or thoughts into long-term memory. Additionally, measuring the relative patterns of left-hemisphere (approach) and right-hemisphere (withdrawal) activation in the pre-frontal area of the brain can assess the motivating potential of the commercial. Like behavioral "persuasion" pre-testing (Rossiter and Eagleson, 1994), brain imaging does not rely on self-reports.
There is no doubt that brain imaging provides the richest diagnostic method for evaluating TV commercials, and the fast recording permitted by SSPT brain imaging can locate strengths and weaknesses in the commercial with very high accuracy. Brainimaging pre-tests can be conducted at reasonable cost with quite small samples of consumers.

Brain imaging is the newest of the psychophysiological methods that appear likely to become the new wave of advertising pre-testing. Another psychophysiological method that has seen great technical advance is eye-tracking, which is especially useful for pre-testing print advertisements (Pieters, Rosbergen, and Wedel, 1999). For pre-testing TV commercials, we think that SSPT brain-image research will be more informative and precise than EEG research (Rothschild and Hyun, 1990), modern electrodermal response (EDR) research (LaBarbera and Tucciarone, 1995), and electromyographic response (EMG) research (Hazlett and Hazlett, 1999). This recent development, we believe, brings us into an exciting new era of insight into how advertising works.

**Appendix: Technical Details of the SSPT Apparatus, Recording, and Data**

Participants are fitted with a lightweight helmet containing multiple electrodes on the scalp to measure cortical electrical activity, which is passed to the SSPT recorder. The helmet has a visor in which is induced a visual flicker stimulus (13 Hz of flickering white light) that evokes a baseline electrical response in the cortex. It is against this baseline that speed of response to the advertising stimuli is recorded. The actual brain electrical activity measure is the latency, the inverse of which is speed, of the steady-state visually evoked potential, SSVEP, relative to the baseline response elicited by the irrelevant flicker. The latency of the SSVEP response indicates how quickly or slowly the brain is reacting to the incoming stimuli.

In this study, the brain electrical activity is recorded from the left and right hemisphere frontal sites thought to participate in the formation of long-term memory traces, namely C3-F7 on the left hemisphere and C4-F8 on the right hemisphere (Buckner, Kelley, and Peterson, 1999). Activity from six other sites is also recorded for other purposes and is not analyzed here. As is usual in cortical measurement, the average electrical potential of both earlobes serves as a reference recording and a nasal electrode serves as a ground. The brain electrical activity is amplified and bandpass-filtered (3 c/sB down at 0.1 Hz and 80 Hz) prior to digitization to 14-bit accuracy at a rate of 250 Hz.

Brain-image readings are taken at the rate of 13 times per second; that is, there are approximately 390 readings during a 30-second TV commercial. To derive the SSVEP response data, smoothed Fourier coefficients are computed, which average 32 readings prior to and 32 readings following the focal reading and give more weight toward the central point. These coefficients are then transformed to polar form (magnitude and phase). The changes in phase are expressed as changes in latency. The data are then averaged across all subjects. The minima of the latencies indicate fastest brain electrical activity responses, while the maxima of the latencies indicate the slowest responses.

Figure 2 illustrates the technical data for one of the test commercials. (See next page.)
Figure 2 Pooled SSVEP Latency Variations at the Left Frontal Site (C3-F7) During One of the 30-Second Television Commercials (Group Data)

Subsequent recognition scores of two frames extracted from the commercial coinciding with latency maxima (slow responses) and latency minima (fast responses) are indicated on the Figure. The horizontal line represents the mean SSVEP latency of all commercials viewed.

REFERENCES


