Journal of Advertising Research
Vol. 51, No. 4, 2011

www.journalofadvertisingresearch.com

Which Broadcast Medium Better Drives Engagement? Measuring the Powers of Radio and Television with Electromyography and Skin-Conductance Measurements

James Peacock

Peacock Research, Inc.

Scott Purvis

Gallup & Robinson, Inc.

Richard L. Hazlett

Johns Hopkins University School of Medicine and Gallup & Robinson, Inc.

INTRODUCTION

"Engagement" is an evolving concept in advertising. In recent years initiatives supported by such organizations as the Advertising Research Foundation, the American Association of Advertising Agencies, and the Association of National Advertisers were formed to develop the concept of engagement as a new planning metric, complementing (and possibly replacing) frequency in media plans.

Whether engagement proves to be an actual new standardized measurement for advertising planning and buying, even the idea of engagement for understanding media and advertising performance is an important one. And although the development of a single, standard, and widely accepted engagement definition is likely to be a difficult, time-consuming undertaking, there are a number of meaningful ways to think about engagement in looking at a medium's ability to communicate and affect advertising within it.

One of the major factors thought to contribute to engagement is the emotional activation of the consumer while viewing or listening to advertising. Advertising that is engaging establishes an emotional connection between the consumer and the brand.

Many advertising-measurement techniques—especially those that look at emotion using survey research methods—depend at least in part on the cognitive side of the messaging. Without minimizing the importance of these "self-report" metrics, though, recent work in neurophysiology has suggested that much information processing takes place largely outside of conscious awareness.

Sensory inputs, such as those from commercials, can be transmitted directly to the amygdala, the emotional center of the brain, and/or indirectly to areas in the neocortex, where complex thought occurs. Neuromarketing research has shown that emotional reactions to advertisements and products do not involve areas of the brain associated with conscious thought (Ambler, Ioannides, and Rose, 2000; Hubert and Kenning, 2008). Therefore, verbal responses (such as those in a survey questionnaire) that are cognitive-based can have clear deficiencies in describing emotion-based response.

Emotional stimuli are evaluated pre-consciously and reacted to even before one can think about how one feels. Consumer decision making is based on an interplay of cognitive and emotional structures and circuits in the brain. For example, when rational messages in advertising tap into personal values important to the consumer, emotion areas of the brain are stimulated. The stronger the emotional bond to the brand (or the emotional messaging of the advertisement), the more important the emotional reaction of the consumer becomes in purchasing decisions.

A great deal of advertising research now points to the conclusion that, "an emotional reaction needs to be established before further cognitive processing of an advertising stimulus takes place. Emotions can be considered as the gatekeeper for further advertisement processing" (Poels and Dewitte, 2006, page 33).

As not everything that consumers are feeling about an advertisement or brand can be expressed in words—or even with pictures—a full understanding of emotional responses to advertising needs to go beyond what a consumer knows how to explain or illustrate. (For an excellent summary of the different methods of measuring emotional responses to advertising, see Poels et al., 2006.)

Charles Darwin first wrote about the changes in human facial expression reflecting an individual's current emotional state and providing a means of communicating emotional information (Darwin, 1872). The face is the best emotional readout humans have. Changes in facial expressions are naturally occurring indicants of positive and negative emotions that are independent of self-report.

A sensitive, precise measure of facial muscle activity is by electromyographic (EMG) technique. This involves placing tiny sensors over specific facial muscles and measuring the electrical current associated with muscle contraction. Facial EMG can detect responses to weakly evocative emotional stimuli even when no change in facial expression is observed (Cacioppo, Petty, Losch, and Kim, 1986) or when subjects are asked to inhibit their facial expression (Cacioppo, Bush, and Tassinary, 1992).

In general, corrugator (frown muscle) EMG activity is positively associated with negative emotional stimuli, negative mood, and increased tension. Zygomatic (smile) EMG activity, conversely, is positively associated with positive emotional stimuli and positive mood states (Lang, Greenwald, Bradley, and Hamm, 1993; Larsen, Norris, and Cacioppo, 2003).

In 1999, facial EMG was first applied to test emotional responses to television commercials (Hazlett and Hazlett, 1999). Two years later, the use of facial EMG for measuring the emotional responses to radio advertisements was validated (Bolls, Lang, and Potter, 2001). Facial EMG also has been used to measure emotional responses to computer interfaces, new products (Hazlett and Benedek, 2007), and Web pages (Hazlett, 2003). Numerous studies have demonstrated that facial EMG is a sensitive and robust measure of emotional responses to media.

For the current study, two EMG measures were taken:

- Positive activation, the zygomatic (smile muscle) EMG
- Negative activation, the corrugator (frown muscle) EMG.

Positive and negative emotional activations are measured separately because they are indicative of separate evaluative processes, which are independent motivators of consumer behavior (Cacioppo, Gardner, and Berntson, 1999). More specifically:

- An increase in positive activations can occur with warm positive feelings, or congruency between the message and personal value of the consumer, and also a sharp positive spike is associated with humor.
- Negative activation can occur when the viewer experiences a negative emotion, which might be an intent of the execution in a negative advertisement. Also, the negative activation can increase gradually during the buildup of anxiety and tension associated with drama; that is then followed by a moderate-to-large increase in positive when the story reaches its climax and there is a positive resolution. Negative activation also can be an unintended reaction to an element of the commercial.

For this study, the authors also included (for an additional baseline) a more traditional arousal measure based on skin conductance. Combined, EMG and skin conductance provide two independent indicators of the viewer's reaction. EMG offers an indication of the positive or negative direction of the emotion (or valence), whereas the skin conductance data provides an indicator of arousal.

Understanding how the aural appeals of radio work relative to the visual/aural appeals of television is an important and new avenue of inquiry. There is strong interest in understanding how advertisements in one medium might affect emotions differently than advertisements in another medium. The existing studies validating the facial EMG method gave the authors confidence that this technique could help further this understanding.

ADVERTISEMENT SELECTION AND PRE-TESTING

To select the advertisements used in the study, the authors worked with a third-party advertising-monitoring service to find pairs of radio and television advertisements that had aired in the last year and that approximately:

- had similar themes for the identical products,
- were of reasonably high audio and video quality, and
- were of standard advertising lengths.

The resulting set of 24 paired commercials then were copy-tested using a standardized online copy-testing service and then assessed against each other and against normative data. That allowed the authors to narrow the pairs of advertisements to the final set of 16 such that the radio/television pairs were the most likely to be "fair" comparisons (radio versus television) based on comparable pre-testing results and to be congruous in theme and content. Those pretesting measures included recall, persuasion, brand rating, likability, and purchase intent.

The final 16 advertisements covered a range of product categories:

- Auto/Car Brand (two campaigns)
- Beverages
- Communications
- Discount department store
- Fast food
- Financial/Investing
- Grocery/canned
- Grocery/deli
- Insurance
- OTC/headache
- OTC/hygiene
- Portable electronics
- Public service organization
- Restaurant
- Travel/resort.

THE PROGRAMMING CONTEXT

A key design component of this study was embedding advertisements within programming. This was not a "copy test" per se; it was a test of advertisements *in context*. The authors chose this study design because consumer responses to commercials that are embedded in programming would be more realistic and valid than responses to an unnatural stand-alone viewing.

In an effort to provide a reasonable cross-section of choices, the authors selected five actual recent radio programs and five actual television programs. Although they did not assert that any given respondent would have found his or her "favorite" on our list, the authors' aim was to provide content types that offered a positive environment for each respondent so that they could fairly assess advertising

engagement in a reasonably engaging programming context.

The types of radio programs that were offered to the survey sample:

- News/Talk (feature-type)
- Soft rock
- Country music
- Urban music
- Classic rock.

And for television, the authors offered the following types of program, purposely choosing programs from less-watched channels in hopes that these shows were unlikely to have been viewed before by the respondents:

- Hard-news documentary
- Soft-news biography
- Female-oriented drama
- African-American-oriented comedy
- General-audience comedy/drama.

METHODOLOGY

Respondents were pre-recruited by telephone from a national field-service list. These samples are built using a variety of opt-in sources, including advertising, the Internet, word of mouth, and mall. Potential participants were selected randomly from the list in two cities and then screened to meet specific age, ethnicity, and experiment-related criteria. Any respondent who had participated in a mall study of any kind within the past year was screened out.

After being told the nature of the measurement and the purpose of the study—the evaluation of either television or radio programming, depending on the pre-assigned group—the respondents were invited to a central facility in one of two major markets (Baltimore and Chicago) for a 30-minute interview. The Baltimore site was a standard focus-group market-research facility located in an office complex in a typical suburban neighborhood about 3 miles from the downtown area. The Chicago site was a mall market-research facility located in a suburban neighborhood approximately 10 miles from the downtown area.

Cooperation fees were offered to increase participation rates. Additional appointments were scheduled after no-shows to ensure that the target number of completes and demographic targets were achieved.

The sample consisted of men and women ages 18 to 54 who used television or radio at least two hours per week. The sample size was 80 each for the radio and television groups. Although this sample size is considered small in the context of attitudinal measurement research, it is large and robust for physiological measurement.

The lab setting was designed to simulate a living room with comfortable furniture. Respondents were tested one at a time. They were hooked up to the measurement equipment and asked to listen to (or watch) a 15-minute sequence of programming material and commercials. This forced exposure method is a well-used one in advertising research, supported by numerous published studies. Additionally, respondents were able to select programming material of interest to them. As the material was listened to or viewed, continuous EMG and other activation measures were recorded (after a brief period of “settling in” to establish a baseline level) using generally accepted skin-preparation and electrodes-placement guidelines (Tassinary, Cacioppo, and Geen, 1989).

For each medium (television or radio), two pods of four commercials each were embedded in the programming material. Thus, a total of 16 pairs of radio and television commercials were tested, with each respondent having been exposed to eight of those commercials for one medium. The order of commercials was rotated such that each commercial had an opportunity to occupy various pod positions.

THE ANALYSIS APPROACH

To test the comparison of the effects of radio and television advertising, the authors examined the following:

- For radio and television advertising overall:
- Mean positive and negative EMG (emotional) activation levels
- Mean overall arousal levels (skin conductance)
- Mean brand recall levels.
- For the 16 individual pairs of advertising campaigns, by television and radio:
- Mean positive EMG (emotional) activation levels
- Mean overall arousal levels (skin conductance).

In the sections that follow, the authors present a few measures that require some definition:

- EMG scores (in general): For the EMG measurement, respondents were given a few minutes at the start of each session to settle in (and settle down) before the program material began. That allowed for the generation of baseline measures on the physiological data. In all EMG charts that follow, that baseline is set to equal a score of 100. Thus, a “positive EMG score” of 117 for the advertising represents an EMG reading that was 17 percent greater than the pre-exposure reading of 100.
- Positive EMG: When the authors present positive EMG scores, those are the readings taken from the “smile muscles” or, more properly, the zygomatic muscles. These are indications of positive emotional responses.
- Negative EMG: When the authors present negative EMG scores, those are the readings taken from the “frown muscles” or, more properly, the corrugator brow frown muscles. These are indications of negative emotional responses and level of tension.
- Arousal scores: These represent the overall levels of skin conductance readings during the advertisements and are indicators of sympathetic nervous system activity, or arousal. Because skin conductance measures are highly variable across respondents, it is common practice to exclude “outliers” and to use the log of the scores.

RESULTS AND DISCUSSION

The authors conducted a (2 × 3) multivariate analysis of variance on the mean scores of the 32 commercials with media as the bi-level factor and the three physiological measures of:

- mean positive emotion: positive EMG scores;
- mean negative emotion: negative EMG scores; and
- mean arousal score: skin conductance log.

As the participants were counterbalanced across the different conditions, gender, age, and product-category effects did not need to be controlled for statistically because the authors controlled for such considerations in their experimental design. With a total participant number of 160, the authors were able to carefully balance their groups and conditions. With counterbalancing, they also were able to control for programming effects on the responses to commercials. In past studies, the authors have found that facial EMG was robust enough to pick up commercial differences in this embedded-in-programming design.

The overall multivariate analysis of variance indicated there were significant differences for the media factor (F (3,28) = 2.80, p = 0.05).

In conducting follow-up analysis of variance (ANOVA) on each measure, the authors found for the positive emotion measure not a significant difference between television and radio (F (1,30) = 0.10, p = 0.75).

There also was not a significant difference found between television and radio on the skin conductance measure of arousal (F (1,30) = 2.25, p = 0.14).

There was, however, a significant difference found between television and radio for the negative emotion factor (F (1,30) = 4.72, p = 0.03). These results indicate that radio and television demonstrated a similar positive emotional impact and arousal level (See Figures 1 and 2), but television commercials had an overall greater negative activation level than the radio commercials. The difference may seem slight, but the methods and measure had enough sensitivity and power to find the difference significant (See Figure 1).

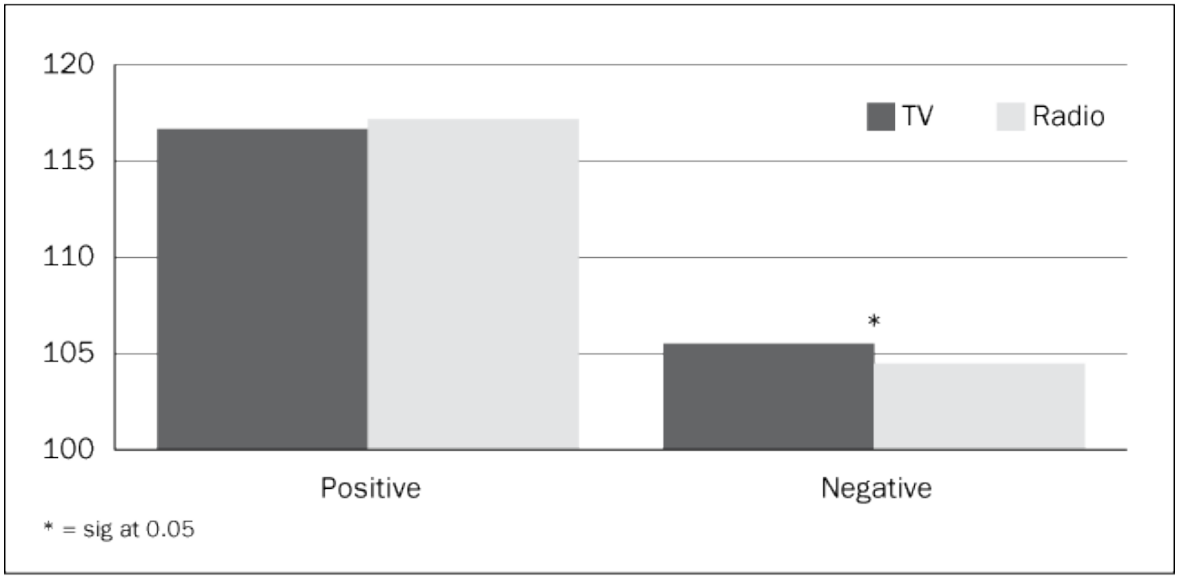

Figure 1 Average EMG Scores

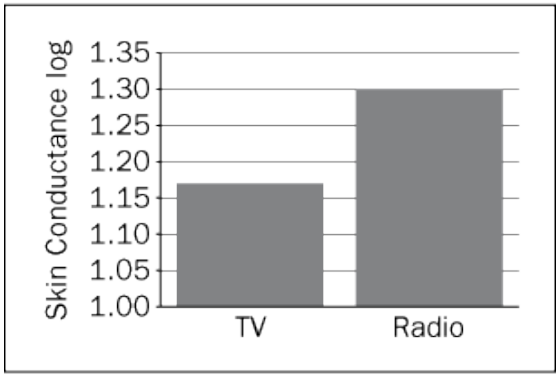

Figure 2 Average Arousal

Some people may argue that achieving emotional impact is not inherently useful unless it can be shown to have another benefit to the advertiser. One way to examine that linkage is to see how these results compare to more traditional and more used measures of advertising effectiveness.

The authors recognize that commercials have differing communication objectives and that no one measure can adequately assess the business value of every commercial. The authors also recognize that brand recall is both a controversial and incomplete measure of "advertiser benefit." Across these 16 campaigns, however—each with its own different purpose—the authors think that stopping power and brand-name registration would have been a plausible part of their communication platforms. So they have used advertising recall as their business-value measure because it could be applied consistently and meaningfully across all these disparate marketing efforts.

There was no statistically significant difference in recall between the radio and television advertisements (F (1,30) = 0.33, p = 0.57). The percentage of participants that could recall the advertised brand after these embedded exposures was essentially the same for radio and television (See Figure 3).

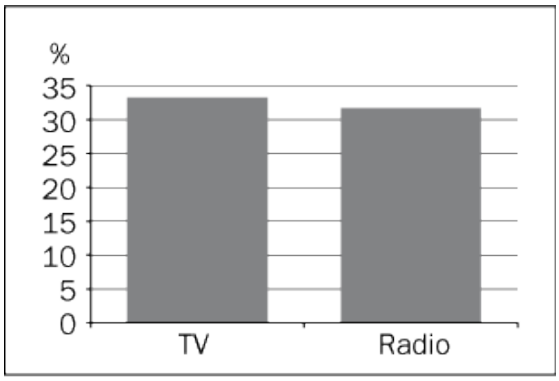

Figure 3 Brand Recall

VARIATIONS BY AD CAMPAIGN

Among the 16 different pairs of radio and television ads, the authors did see several in which radio delivered stronger positive emotional impact than television and one in which the reverse was true (Figure 4).

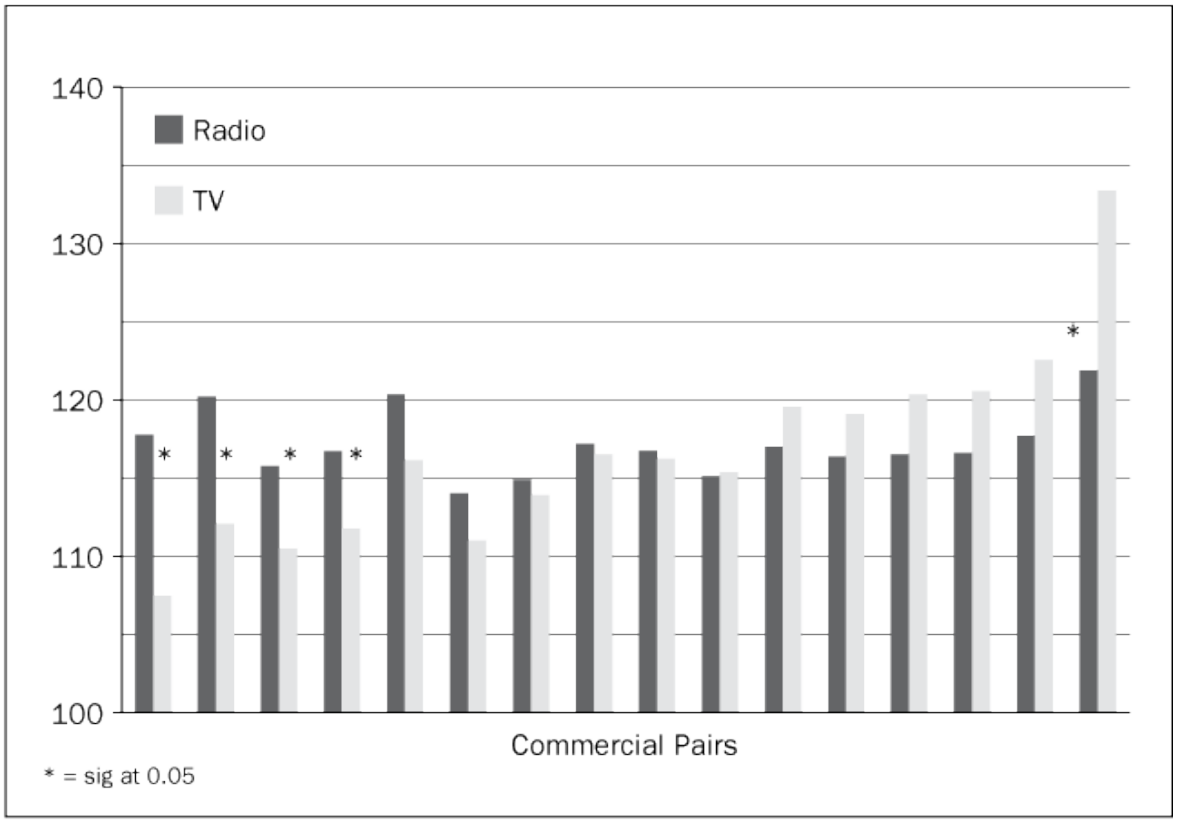

Figure 4 Positive Activation

With 80 subjects viewing or listening to one member of each commercial pair, the authors conducted a two-way ANOVA for each commercial pair with media as their two-level factor. Five of the 16 pairs were significantly different on the media factor at (at least) the 0.05 level. The ANOVA statistic for the strongest radio commercial in the first position was (F (1,77) = 8.45, p = 0.005), and the significantly greater positive television commercial in the last position was (F (1,77) = 5.38, p = 0.023; See Figure 4).

Negative-activation pair differences on the media factor also were analyzed with the same ANOVA approach used with the positive-activation scores. This analysis found two commercial pairs [(F (1,77) = 4.61, p = 0.035) and (F (1,77) = 9.58, p = 0.003)] with significantly greater negative activation for the television pair member.

Arousal level—as measured by skin conductance log—also was ANOVA-tested for pair differences on the media factor. This analysis found three commercial pairs significantly different at the 0.05 level on the media factor. One of the three significant analyses was for the same commercial pair found significant for positive activation; one for the negative (with the same pair member congruently elevated); and one did not have significant valence differences.

These results reflect what is known about skin conductance findings generally. Elevated arousal levels can be associated with either elevated positive or negative emotion and also without either being elevated, and the arousal is not reflective of a complete emotional response.

The authors also examined the relationship between the emotion measures and brand recall by medium for the commercial pairs. Brand recall was the percentage of participants who recalled the commercial. For radio only, the correlation between positive activation and brand recall was significant, and that highly so (Table 1). The correlation between positive activation and brand recall for television was still substantial and just missed being significant (p = 0.06). As the emotion measures were collected during the commercial and recall after, the authors were able to conduct a regression analysis with recall as the dependent variable, and found that positive activation explained almost half of the variance in recall for radio (R^2 = 0.48, p < 0.01), and more than a fifth for television (R^2 = 0.22, p < 0.06).

TABLE 1

Correlation between Emotional Measures and Recall by Medium

	Positive	Negative	Arousal
Radio			
Recall	0.69**	–0.01	0.34
Television			
Recall	0.46	0.13	0.14

*Notes: ** Significant at 0.01; n = 16*

CONCLUSIONS

This study was designed to assess how well radio advertisements can generate emotional responses and engage with consumers as compared to television ads. Advanced methods were used that measure emotional activation without requiring a verbal response. The facial EMG methods demonstrated a sensitivity to finding differences in emotional response between commercials and type of media.

After evaluating 16 different real advertising campaigns within actual programming, it appears that:

- Radio and television communicate positive emotional messages about equally. In spite of just being an audio medium, radio advertisements had a positive emotional impact on consumers that was at a similar level of engagement as television ads.
- Television may evoke a slightly greater negative emotional reaction in some commercials than similar radio ads. What is intriguing is why the greater elevation for television, particularly as negative activation is not generally very elevated for commercials except in negative advertising campaigns.
- Perhaps the narrower range and smaller variance of negative activation allow for greater sensitivity in detecting subtle differences. Future studies that examine content and execution strategy will need to be conducted to better understand this slight elevation in negative activation for television.

The embedded commercial-in-programming design of this study allowed the authors to take a valid measure of brand recall. Positive emotion and brand recall were found to be positively correlated. And, because the emotion measures used in this study were collected during the commercial and brand recall after, the authors found a directional relationship leading from the consumer's positive emotional experience during the commercial to recalling the brand at a later time.

One of the advantages of this study was that positive emotion, negative emotion, and arousal were measured separately. With separate measurements, the authors were able to see that there were differences in how each aspect of emotional response was affected by medium type and related to recall.

The fact that positive emotion was more strongly related to recall than negative emotion or arousal demonstrates that it is important to consider emotional valence when addressing the relationship of emotion to advertising effectiveness.

This study begins an investigation into the similarity and differences in how radio and television may emotionally affect the consumer and communicate the brand message. The authors acknowledge that this single testing environment is a limitation with unknown effects. For example:

- A significant amount of radio listening occurs in cars, and the authors' testing environment could not simulate that listening condition.
- A significant amount of radio listening and television viewing takes place in a social context. The authors do not know whether that or other environments would yield better or worse results for radio and/or television advertising.

These were necessary limitations, though, to directly compare radio and television advertising in the context of programming. To answer these additional questions, future studies will need to be designed to compare the effects of living room and car listening environments or varying levels of social involvement.

References

Ambler, T. A. Ioannides, and S. Rose. "Brands on the Brain: Neuro-Images of Advertising." *Business Strategy Review* 11, 3 (2000): 17–30.

Bolls, P. D., A. Lang, and R. F. Potter. "The Effects of Message Valence and Listener Arousal on Attention, Memory, and Facial Muscular Responses to Radio Advertisements." *Communication Research* 28 (2001): 627–651.

Cacioppo, J. T., L. K. Bush, and L. G. Tassinary. "Microexpressive Facial Actions as a Function of Affective Stimuli: Replication and Extension." *Psychological Science* 18 (1992): 515–526.

Cacioppo, J. T., W. Gardner, and G. Berntson. "The Affect System Has Parallel and Integrative Processing Components: Form Follows Function." *Journal of Personality and Social Psychology* 76 (1999): 839–855.

Cacioppo, J. T., R. E. Petty, M. E. Losch, and H. S. Kim. "Electromyographic Activity over Facial Muscle Regions Can Differentiate the Valence and Intensity of Affective Reactions." *Journal of Personality and Social Psychology* 50 (1986): 260–268.

Darwin, C. The Expression of Emotions in Man and Animals. London, UK: John Murray, 1872.

Hazlett, R. L. "Measurement of User Frustration: A Biologic Approach." In Proceedings of CHI 2003 Conference on Human Factors in Computing Systems. Ft. Lauderdale, FL: ACM Press, 2003.

Hazlett, R. L., and J. Benedek. "Measuring Emotional Valence to Understand the User's Experience of Software." *International Journal of Human-Computer Studies* 65 (2007): 306–314 [Special Issue on Evaluating Affective Interactive Systems].

Hazlett, R. L., and S. Y. Hazlett. "Emotional Response to Television Commercials: Facial EMG vs. Self-report." *Journal of Advertising Research* 39, 2 (1999): 7–23.

Hubert, M., and P. Kenning. "A Current Overview of Consumer Neuroscience." *Journal of Consumer Behaviour* 7 (2008): 272–292.

Lang, P. J., M. K. Greenwald, M. M. Bradley, and A. O. Hamm. "Looking at Pictures: Affective, Facial, Visceral, and Behavioral Reactions." *Psychophysiology* 30 (1993): 261–273.

Larsen, J. T., C. J. Norris, and J. T. Cacioppo. "Effects of Positive and Negative Affect on Electromyographic Activity over Zygomaticus Major and Corrugator Supercilii." *Psychophysiology* 40 (2003): 776–785.

Poels, K., and S. Dewitte. "How to Capture the Heart? Reviewing 20 years of Emotion Measurement in Advertising." *Journal of Advertising Research* 46, 1 (2006): 18–37.

Radio Ad Effectiveness Lab Inc. (2005). "Radio's ROI Advantage." Retrieved from http://RadioAdLab.org.

Radio Ad Effectiveness Lab Inc. (2006). "Personal Relevance Two: Radio's Receptive Ad Environment." Retrieved from http://RadioAdLab.org.

Tassinary, L. G., J. T. Cacioppo, and T. R. Geen. "A Psychometric Study of Surface Electrode Placements for Facial Electromyographic Recording: I. The Brow and Cheek Muscle Regions." *Psychophysiology* 26 (1989): 1–16.

About the authors

James Peacock is the president of Peacock Research, Incorporated. Among other clients, Jim serves as technical consultant to the Media Rating Council; he also served as research consultant to the Radio Ad Effectiveness Lab. Prior to forming Peacock Research, he was VP/Research at Arbitron, leading the departments that were responsible for all television and radio methodology research. He has written and spoken extensively for media research industry conferences including The Advertising Research Foundation, ESOMAR, the RAB, NAB, BBM Canada, IBOPE Latin America, the Media Research Council of Chicago, and others. Email: jpeacock@peacockresearch.com

Scott Purvis is president of Gallup & Robinson, Inc. With 25 years of experience in advertising research, he works directly with leading companies in packaged goods, technology, pharmaceuticals, automotive, and financial services to help them measure, understand, and

improve the effectiveness of their advertising. He is also the author of Which Ad Pulled Best? now in its tenth edition, and a co-developer of CERA, a physiological technique for measuring emotions-based response to advertising. Scott holds an undergraduate degree in linguistics, a JD from Georgetown University, and an MBA from George Washington University. Email: sos@gallup-robinson.com

Richard Hazlett is senior scientist at Gallup and Robinson and also holds a part-time position as assistant professor in the department of psychiatry and behavioral sciences at Johns Hopkins University School of Medicine. Dr. Hazlett has presented his findings and methods at numerous conferences and in publications in leading journals. Richard Hazlett has a PhD in clinical psychology from Illinois Institute of Technology in Chicago and, before starting his faculty position in 1992 at Johns Hopkins, he completed a postdoctoral fellowship there in psycho-physiological and emotion research. Email: rlhazlet@jhmi.edu