

# Comparative analysis of emotion measures

## A case study using the Heineken "Weasel" commercial

In recent years, interest in studying the role of emotions in consumer decisions has spiked significantly. This is consistent with the increased focus on emotion in many areas of psychology and economics. The complex nature of the emotional process, which involves physiological and behavioral changes, has led to the development of new constructs for measuring emotions-based response and the enhancement of longstanding ones that go well beyond simple question-and-answer surveys. These systems have underlying similarities (e.g., they are generally nonverbal and provide continuous or nearly continuous measurement) but also underlying differences (e.g., they measure different aspects of emotional activation and have different analytic protocols). We will look in detail at three such systems: Picture Sorts (PS), as executed by Ameritest, Albuquerque, N.M.; the Biometric Monitoring System (BMS), as executed by Boston-based Innerscope; and facial electromyography (FEMG), as executed by our firm's Continuous Emotional Reaction Analysis (CERA).

### Three main protocols

Researchers studying emotions have relied on three main protocols to study emotions: self-reports (descriptions of feelings); physiological measurements (measures of blood pressure, heart rate, brain activity, etc.); and behavioral/expressive measures (eye, face and vocal expressions).

### Self-reports

In this protocol, researchers simply ask people to describe their current, past or typical emotions. The basic premise is that people know their own emotional states the best. Self-reports are straightforward data to

collect and enable researchers to tap into the emotional fabric, especially when they are interested in observing changes in emotion over time. But data from self-reports are not nearly as straightforward to interpret as they seem. This is because they carry with them numerous socio-psychological factors that can influence the validity of a response at generally unknown levels. These include the ability of a respondent to articulate his or her deeper thoughts or feelings; the motivation of the respondent to respond fully; the urge to respond in socially desirable ways; and biases intro-



By Sandeep Patnaik  
and Scott Purvis

### snapshot

A number of different approaches have emerged for measuring emotional response to advertisements. This article compares and contrasts the results of three different measurement techniques as they were applied to the same 30-second TV spot for Heineken beer.

*Editor's note: Sandeep Patnaik is research director at Gallup & Robinson Inc., a Pennington, N.J., research firm. He can be reached at 609-730-1550 or at sandeep.patnaik@gallup-robinson.com. Scott Purvis is president of Gallup & Robinson Inc. He can be reached at 609-730-1550 or at sos@gallup-robinson.com. To view this article online, enter article ID 20110307 at [quirks.com/articles](http://quirks.com/articles).*

duced by the construction of the questioning sequences, etc. Also, the precision of self-reports often suffers because they are retrospective and because each person's standard of comparison differs.

### **Physiological measurements**

Several theorists, including Damasio (1994), argue that thought and emotions cannot occur independent of the body. When we think emotionally, our whole body is readied for action and we receive feedback from it. Consequently, a number of studies have turned to physiological measures to tap activity in the nervous systems, i.e., autonomic nervous system (ANS) and the central nervous system (CNS), to understand the dynamic interplay of cognition, arousal and emotion.

#### ***Autonomic nervous system:***

Many emotional conditions are states of intense arousal that are directly related to the activity of the autonomic nervous system, which has two branches, viz., the sympathetic nervous system and the parasympathetic nervous system. The sympathetic nervous system (SNS) is related to increased heart rate, breathing rate, sweating and adrenaline secretions. The parasympathetic nervous system (PNS) conserves energy to use in growth and development. Most studies of emotion physiology have focused on the SNS rather than PNS. However, this emphasis of the SNS favors the study of negative emotions. A positive emotion, like happiness or love, that is generally associated with less intense SNS activity may not be as arousing as a stronger negative emotion like anger or fear.

#### ***Central nervous system:***

Measuring brain activity to gauge emotional response gained increasing popularity in the 1990s. One of the most common techniques to measure brain activity is electroencephalography (EEG), which registers variations in brain waves produced by the cortex. A major strength of an EEG is that researchers can link electrical changes in general brain areas to exposure to emotional stimuli. However,

EEGs record activities of brain cells nearest to the electrodes and cannot access information from the deep brain areas that are especially important for emotion.

#### ***Behavioral/expressive measures:***

Apart from physiological indices, emotion researchers have also explored behavioral and expressive aspects of emotions.

The face is the most expressive part of the body and has been the target of attention by most researchers. Two major approaches to facial measurement include the facial coding schemes and electrophysiological recordings.

Ekman and Friesen's (1978) facial action coding system (FACS) is perhaps the best known among the facial coding schemes. In this approach, trained coders manually code facial expressions by identifying which muscles are contracted on a person's face at any given moment and recording how intensely and how long those muscles contract. Researchers use the patterns of muscle contraction as a nonverbal measure of people's emotions. FACS is quite useful for researchers studying the effects of emotion on social interaction. However, coding of facial expressions is very time-intensive and is greatly dependent on the coder's ability to consistently distinguish emotional movements from all of the other facial movements.

Another way of measuring facial expressions of emotions is through facial electromyography (FEMG). This technique measures electrical potentials from two major muscle groups in the face, the corrugator supercilii and zygomaticus major, via the placement of surface electrodes on the skin of the face. It is seen that activity of the zygomatic muscle, which controls smiling, is associated with positive emotional stimuli and positive mood state. In contrast, activity of the corrugator muscle, which lowers the eyebrow and is involved in producing frowns, varies inversely with the emotional valence of presented stimuli.

Researchers have been successful in adapting FEMG as a method to study emotional expression to advertising stimuli (Hazlett, 1999).

Bolls et al. (2001) found that zygomatic muscle activity was stronger during radio advertisements with a positive emotional tone whereas corrugator muscle activity was greater during ads with a negative emotional tone. FEMG has been shown to be capable of measuring response to weakly evocative emotional stimuli even when no changes in facial displays have been observed with the FACS system as well as when subjects were instructed to inhibit their emotional expression.

It is possible to broadly relate the three measurement systems at which we will be looking - PS, BMS and CERA - to the above emotions-research protocols - self-report, physiological and behavioral/expressive. PS emphasizes analysis of subjective experiences reported by the respondents. BMS relies primarily on physiological data gathered from respiratory, heart and motor responses. CERA is based on the behavioral/expressive FEMG technique.

### **Significantly different**

All three systems are significantly different from each other in their methodology as well as data analysis.

Picture Sorts uses visual sorting of photographic stills taken from the commercial to probe respondent reactions shortly after initial viewing, based on the respondent's memory, wherein the respondent attempts to reconstruct the experience after the initial exposure by use of picture-sorting cards. Multiple sorts are used to obtain multiple levels of self-reported response. Typically the respondents are expected to sort images based on those they recognize seeing.

The Biometric Monitoring System is a physiological measure that involves embedding biometric sensors in a vest or garment worn by the respondent. The embedded sensors detect and record biological measures such as respiration, heart rate, skin conductance, etc.

Continuous Emotional Response Analysis utilizes facial electromyography techniques to measure emotional responses to advertise-

Table 1: Comparison of Samples, Mode of Exposure and Method of Study

	Facial EMG	Picture Sorts	Biometrics
Sample size	40 males; 21-35 years of age.	203 males; 21-35 years of age.	30 males; 25-35 years of age.
Mode of exposure	Embedded in a 15-minute video	Online clutter exposure	Embedded in 30-minute TV program
Method of study	FEMG; measurement of contraction of facial muscles	Visual picture sorts of key executional frames; respondents interviewed online	Biometric Monitoring System; sensors embedded in garment measuring respiration, motion, heart rate and skin conductance.
System (company providing it)	CERA (Gallup & Robinson)	PS (Ameritest)	BMS (Innerscope)

Table 2: Comparative Analysis of CERA, BMS and PS with Reference to Selected Key Events

Data Points	Descriptions of Event	Facial Electromyography (CERA)	Biometric Monitoring System (BMS)	Picture Sorts (PS)
A	Opening scene showing a man walking on the street, holding a brown bag.	No significant activation.	Flat and steady; little or no activation.	Slight rise in activation.
B	Scene of a party in progress; women in the hallway.	Sharp increase in both positive and negative emotion activation.	Flat and steady; no major positive activation.	Negative trend in activation.
C	Fuzzy shot of people crowding the hallway.	Slight decline in positive activation; negative activation remains high.	No change in either positive or negative activation.	Major decrease in positive activation.
D	Protagonist greeting others guests.	Significant spike in positive activation.	Steady increase in activation.	Sharp increase in positive activation.
E	Close-up of protagonist's (smiling) face.	Sharp decline in positive activation.	Slow and steady rise in positive activation.	Significant drop in positive activation.
F	Exchanges smiles with woman.	Slight increase in positive activation.	Slight increase in activation.	Sharp increase in positive activation.
G	Walks over and opens the fridge.	Modest increase in positive activation accompanied by some decrease in negative activation	Gradual and steady increase in positive activation.	Sharp decline in positive activation.
H	Advertised brand is revealed, "The Weasel" superimposed on the screen.	Positive activation declines along with increase in negative activation.	Sharp increase in positive activation.	Positive activation level reaches its peak.
I	Product switch is made.	Positive activation peaks as humor in the episode is recognized.	Steady decline in negative activation.	Positive activation holds steady.
J	Protagonist rejoins party.	Both positive and negative activation decrease after transition.	Positive activation continues to decline.	Sharp decline in positive activation
K	Dark screen followed by tagline, "It's all about the beer."	Positive stays high at branding; slight decline in negative activation.	Sharp decline in activation.	Positive activation holds steady.
L	Final branding moment.	High activation at payoff.	No payoff at branding moment.	Positive activation steady during final moments.

ments. It relies on the precept that facial muscle movements offer the best "markers" to accompany emotional responses to a commercial. As mentioned previously, the technique involves measuring minute electrical impulses in two major muscle groups in the face, the corrugator supercilii and zygomaticus major muscle groups, which have been shown to be valid indicators

of negative and positive emotional responses, respectively.

In addition to their essentially non-verbal nature, an underlying commonality in all three approaches has been that they yield continuous measures. As such, it is feasible to compare them on a common continuum and highlight areas where their results are either similar or dissimilar.

### "Weasel" television ad

The common stimulus for all three measures was a 30-second Heineken beer television ad titled "Weasel." The commercial features a young man with a winning smile and oozing with confidence arriving at a party carrying a brown bag, probably containing a six-pack of beer. As he heads for the refrigerator, he exchanges looks and smiles with an attractive woman who appears to be checking him out. While stashing his "inferior" brand of beer in the fridge, the man notices another six-pack of a superior beer, i.e., Heineken. As the man grabs a couple of Heineken bottles, a title is superimposed on the screen: "The Weasel." After making the deceptive switch, he rejoins the party. The concluding title reads, "It's all about the beer."

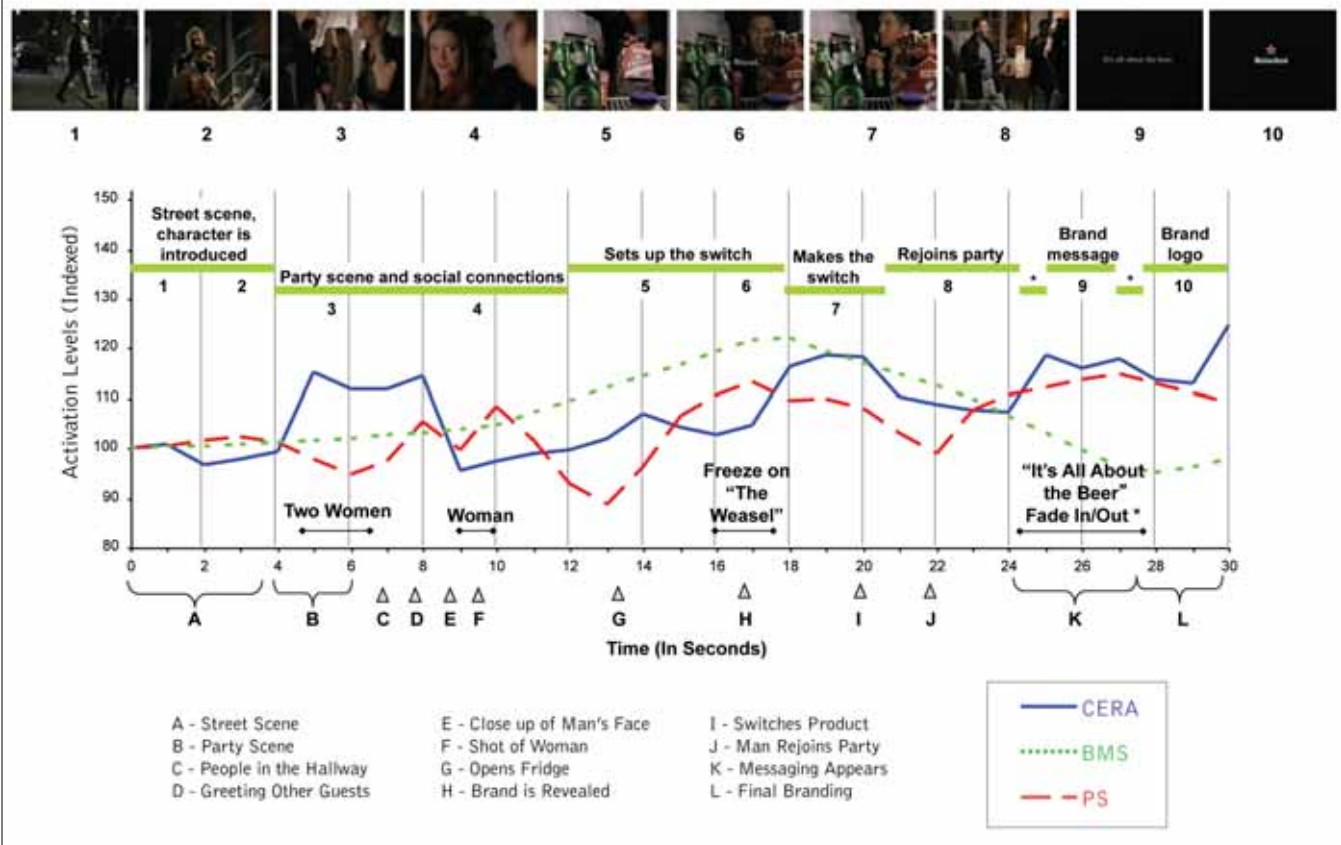
The commercial was tested, the data gathered and the results reported by each company as part of the Emotions in Advertising Project from the American Association of Advertising Agencies (AAAA) and the Advertising Research Foundation (ARF) (Table 1).

A note on the comparative analysis: This analysis is based on findings available in the public domain, mostly as a result of the companies having participated in the AAAA/ARF study. The authors did not have direct access to the data sets obtained by either Ameritest or Innerscope. Comparing three different techniques using totally different methodologies and measurement metrics presented two challenges:

1. While both CERA and BMS are based on "moment-by-moment" physiological reactions on a temporal continuum, PS uses photographic stills of important scenes from the video to probe respondent response. In order to ensure comparability of results, it was necessary to first estimate the time sequence of each of the PS photographs so as to facilitate subsequent mapping of activation levels, measured by each technique against a common (temporal) continuum.
2. Because each of these three techniques used different scales (with different measuring units) to mea-



Figure 1: A Graphical Representation of the Responses Recorded by CERA, BMS and PS



sure the activation levels, it was necessary to index and position all three sets of scores at the same starting point in order to facilitate comparison. The indexed score for each technique was calculated by using the following formula: indexed score =  $(X-SP)/SD$ , where  $X$  = each observation,  $SP$  = starting point for technique, and  $SD$  = standard deviation for each technique. We have used a common term, activation levels (indexed), to represent respondents' reactions in the three techniques.

### Congruency and divergence

Analysis of the results obtained by each of the techniques reveals areas of congruency and areas of divergence. During the opening scene showing the visual of a person carrying a brown paper bag walking toward a house (seconds 0-4), the response level in each of the measures is consistent with little activation being evident; there is nothing remarkable about either the scene or the protagonist.

Beginning at :04 and continuing

to :08, the systems show significant differences. This time coincides with the screen revealing a party scene with young men and women standing in the hallway, chatting and enjoying themselves. PS shows the scene to be initially deactivating; BMS as generally neutral; and CERA as activating. Both PS and CERA show :08, where the actor introduces himself, to be an activating image.

CERA and PS both show sharp decline in audience interest between :08 and :09 while BMS response level remains unchanged. At this point the camera remains focused on one side of the protagonist's face (in contrast to the preceding party scenes).

The appearance of an attractive girl at :09 piques viewers' attention. All the measures record an increase in activation level of viewers, with PS recording the sharpest spike; both BMS and CERA register more modest gains.

When the protagonist walks to the refrigerator with the brown bag (around :10), PS records a sharp decline in activation that continues through :14, when the refrigerator

door opens to reveal the contents. Both CERA and BMS register a steady and growing level of emotional activation during this time.

A key moment in the ad is at :14 where the "inferior" beer brought by the man is juxtaposed with the "superior" Heineken brand. All three measures record increasing activation. However, the actual apex in activation occurs at somewhat different parts of the exchange: for PS and BMS, it is when the word "weasel" appears, and for CERA, it is when the "exchange" takes place.

At about :20 there is a significant transition when the man walks out to rejoin the party after having switched the brand and flaunting the "superior" Heineken bottle. There is decline in positive and negative activation in all measures.

Beginning at around :25, the commercial fades to black, then says, "It's all about the beer," then fades to black again, followed by the Heineken logo at about :28. This is a key moment in the ad that provides another important contrast in the findings of the three measures:

CERA shows that the high positive emotional activation invoked during the brand “switch” is sustained through the end, when the final positioning line and brand name are shown; PS shows activation during the messaging but not during the branding; and BMS shows declining activation during both the messaging and branding.

Table 2 summarizes the findings of the three different methods at the aforementioned key moments of the commercial.

### Significant differences

In general, all three methods were consistent in identifying the peak activation period in the commercial. However, there were significant differences in what the three revealed about the build-up and selling messages.

BMS results were quite linear compared to both CERA and PS. In BMS the activation rose slowly but steadily till about :18 when both the Heineken and “inferior” brand were placed next to each other. Thereafter, the activation dropped steadily until the final branding moment, when it increased slightly. Like other measures of the sympathetic nervous system, BMS does better in recording strong arousals than when the level of activation is mild to moderate.

In contrast, both PS and CERA showed a dynamic and nonlinear response pattern. In the initial moments of the ad, where there was a transition from the street scene to the party scene, CERA proved to be the most sensitive measure in recording the expected spike in interest. Subsequently, there was a very sharp decline in PS measure, especially at the scene where the man was shown standing before the refrigerator; neither BMS nor CERA reveal such a sharp decline.

Another area of difference between the three methods was seen during the brand message and logo presentation toward the end of the commercial. BMS showed a decline in activation during these sequences while PS and CERA showed it being maintained during the messaging and

CERA showed it being additionally maintained during the branding.

The greater apparent sensitivity of PS and CERA seems to make them more actionable measures than BMS as they reveal how the components of a commercial contribute to its overall activation levels. Some of the less intuitive findings from PS (e.g., the peak at :09.5 showing a typical party scene and the valley at :13.5 exposing the brand) are difficult to explain though critics may point to them as evidence of potential problems typically associated with non-coincidental, cognitively filtered self-reporting systems. Even though CERA’s findings seem to be largely in accord with the storyline of this ad, it is worth noting that FEMG, as a technique, is not free from criticism. Some have pointed to the obtrusive nature of the procedure, which may interfere with any spontaneous, natural reactions. Others have noted the possibility of non-affective processes like mental fatigue, task involvement, speech, etc., tainting the measure of electrical conductance in target muscles.

### Unique opportunity

In conclusion, the study was a unique opportunity to compare and contrast three techniques’ relative effectiveness in tracking audience emotional valence during the course of a single commercial. While all three techniques were successful in tracking significant changes in emotional reactions in the viewer, there were significant differences in the extent to which they were able to provide valence information for each. Comparative studies of this nature will be useful to establish the concurrent validity of different types of measures used to assess advertising effectiveness. |Q

### References

- Bolls, P.D., Lang, A., Potter, R.F. (2001). “The Effects of Message Valence and Listener Arousal on Attention, Memory and Facial Muscular Responses to Radio Advertisements.” *Communication Research*, 28(5), pp. 627-651.
- Christianson, Sven-Ake (Ed). (1992). *The Handbook of Emotion and Memory: Research and Theory*. Hillsdale, N.J., England:

Lawrence Erlbaum Associates, Inc.

Damasio, A.R. (1994). *Descartes’ Error: Emotion, Reason and the Human Brain*. New York: G. P. Putnam.

Dimberg, U.(1990). “Facial Electromyography and Emotional Reactions.” *Psychophysiology*, Vol. 27 (5) pp. 481-494.

Ekman, P., Friesen, W.V. (1978). *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press, Palo Alto, Calif.

Hazlett, R.L., Hazlett, S.Y. (1999). “Emotional Response to Television Commercials: Facial EMG vs. Self Report.” *Journal of Advertising Research*, Vol. 39 (2), pp. 7-23.

Levenson, R.W.(1988). “Emotion and the Autonomic Nervous System: A Prospectus for Research on Autonomic Specificity.” In H. Wagner & A. Manstead (Eds), *Handbook of Social Psychophysiology*. New York: Wiley.

Marci, Carl D. (2006). “A Biologically Based Measure of Emotional Engagement: Context Matters.” *Journal of Advertising Research*. Vol. 46 (4), pp. 381-387.

“New Thoughts on Measuring Emotional Response to Advertising” (2006) Retrieved from AAAA/ARF Web site: [http://www.mrcouncil.org/uploadedarchives/MRC%20-%20Jan\\_20\\_06%20Speaker%20presentation-recd%202-10-06.pdf](http://www.mrcouncil.org/uploadedarchives/MRC%20-%20Jan_20_06%20Speaker%20presentation-recd%202-10-06.pdf).

Nunez, P.L., Srinivasan, R. (2006). *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press.

Ohme, R., Reykowska, D., Wiener, D., Choromonska, A. (2009). “Analysis of the Neurophysiological Reactions to Advertising Stimuli by Means of EEG and Galvanic Skin Response Measures.” *Journal of Neuroscience, Psychology and Economics*. Volume 2 (1) pp. 21-31.

Van Boxtel, A., (2010). “Facial EMG as a Tool for Inferring Affective States.” *Proceedings of Measuring Behavior 2010*, Eindhoven, The Netherlands, August 24-27, 2010.

Van Boxtel, A., Goudswaard, P., Van der Molen, G.M., and Van den Bosch, W.E.J. (1983). “Changes in EMG Power Spectra of Facial and Jaw-elevator Muscles During Fatigue.” *Journal of Applied Physiology*, 54, 51-58.

Waterink, W., and van Boxtel, A. (1994). “Facial and Jaw-Elevator EMG Activity in Relation to Changes in Performance Level During a Sustained Information Processing Task.” *Biological Psychology*. 37, 183-198.

Zaltman, G. (2003), *How Customers Think: Essential Insight into the Mind of the Market*. Harvard: Harvard Business School Press.