VINOD VENKATRAMAN, ANGELIKA DIMOKA, PAUL A. PAVLOU, KHOI VO, WILLIAM HAMPTON, BRYAN BOLLINGER, HAL E. HERSHFIELD, MASAKAZU ISHIHARA, and RUSSELL S. WINER*

In the past decade, there has been a tremendous increase in the use of neurophysiological methods to better understand marketing phenomena among academics and practitioners. However, the value of these methods in predicting advertising success remains underresearched. Using a unique experimental protocol to assess responses to 30-second television ads, the authors capture many measures of advertising effectiveness across six commonly used methods (traditional self-reports, implicit measures, eye tracking, biometrics, electroencephalography, and functional magnetic resonance imaging). These measures have been shown to reliably tap into higher-level constructs commonly used in advertising research: attention, affect, memory, and desirability. Using time-series data on sales and gross rating points, the authors attempt to relate individual-level response to television ads in the lab to the ads' aggregate, market-level elasticities. The authors show that functional magnetic resonance imaging measures explain the most variance in advertising elasticities beyond the baseline traditional measures. Notably, activity in the ventral striatum is the strongest predictor of real-world, market-level response to advertising. The authors discuss the findings and their significant implications for theory, research, and practice.

Keywords: advertising elasticities, neuroscience, biometrics, implicit measures, market response modeling

Online Supplement: http://dx.doi.org/10.1509/jmr.13.0593

Predicting Advertising Success Beyond Traditional Measures: New Insights from Neurophysiological Methods and Market Response Modeling

*Vinod Venkatraman is Assistant Professor of Marketing (e-mail: vinod.venkatraman@temple.edu), Angelika Dimoka is Associate Professor of Marketing (e-mail: dimoka@temple.edu), Paul A. Pavlou is Milton F. Stauffer Professor of Information Technology and Strategy (e-mail: pavlou@temple.edu), Khoi Vo is Senior Research Associate (e-mail: khoi.vo@temple.edu), and William Hampton is Research Assistant (e-mail: william.hampton@temple.edu), Center for Neural Decision Making, Fox School of Business, Temple University. Bryan Bollinger is Assistant Professor of Marketing, Fuqua School of Business, Duke University (e-mail: bryan.bollinger@duke.edu). Hal E. Hershfield is Assistant Professor of Marketing, Anderson School of Management, University of California, Los Angeles (e-mail: hal.hershfield@anderson.ucla.edu). Masakazu Ishihara is Assistant Professor of Marketing (e-mail: mishihar@stern.nyu.edu), and Russell S. Winer is William H. Joyce Professor of Marketing (e-mail: rwiner@stern.nyu.edu), Stern School of Business, New York University. The authors are grateful for funding support from the Advertising Research Foundation (ARF) and the corporate sponsors of this study. They thank Horst Stipp from ARF and Jim Thompson from Temple University for their support and guidance throughout the project, Dave Walker from IPSOS-ASI for assistance with the protocol design for the traditional surveys, Jessica Fleck (Stockton College) for assistance with collection of EEG data, and Rene San Martin (Duke University) for assistance with the EEG analyses.

Companies spend millions of dollars each year striving for advertising success. Advertising spending occurs for pretesting copy alternatives while a campaign is being developed as well as for in-market analyses after the campaign is launched. Many methods have been developed to pretest ads, ranging from self-reported "traditional" measures such as recall, liking, and purchase intent to neurophysiological measures such as functional magnetic resonance imaging (fMRI). Furthermore, sophisticated statistical approaches, referred to as marketing-mix modeling, have also been used to evaluate ex post the impact of advertising spending across multimedia.

Marketing textbooks draw a distinction between rational and emotional advertising (Batra, Myers, and Aaker 1996), wherein the former refers to advertising of factual or conscious information and the latter refers to advertising that targets unconscious and emotional processes. Marketers have relied on various approaches to measure rational processes for decades (e.g., Lucas and Britt 1963). These measures include recognition, recall, liking, and persuasion. Marketers have also attempted to measure unconscious automatic reactions to advertising. As Stewart (1984) notes, popular methods have included several physiological approaches such as pupillary response, heart rate, eye movements, voice pitch analysis, and neuroimaging, all of which are commonly referred to as neurophysiological methods.

The past decade has experienced an explosion of research in neuroscience and the use of multiple neurophysiological methods to study marketing, consumer behavior, and advertising phenomena, broadly referred to as consumer neuroscience or neuromarketing (e.g., Ariely and Berns 2010; Camerer, Loewenstein, and Prelec 2005; Dimoka 2012; Smidts et al. 2014; Venkatraman et al. 2012; Yoon et al. 2012). A new industry has also been built around neuromarketing tools in the past decade.[1] This growth is due to a combination of technological advances in fMRI, electroencephalography (EEG), eye tracking, and other neurophysiological tools with increased accessibility of these methods due to decreased administration costs (Dimoka, Pavlou, and Davis 2011).

Although there has been considerable research in both academia and industry using neurophysiological measures to better understand consumer responses to advertising (Ohme et al. 2010; Stipp and Woodard 2011), it is important to examine whether these measures actually translate into real-life advertising success. The overall goal of this article is to link traditional and neurophysiological measures to actual market responses to advertising in terms of advertising elasticities.[2] Although, to our knowledge, no research has directly linked neurophysiological measures to advertising elasticities, one study has found a relationship between neurophysiological measures and market sales. Specifically, Berns and Moore (2012) use fMRI data to predict music popularity by measuring the brain activity of 27 adolescents when listening to 15-second song clips. The authors use three-year data from Nielsen SoundScan to show a significant link between brain activity (ventral striatum) and sales.

A common question practitioners ask is whether neurophysiological methods are "valuable"—that is, do they contribute anything beyond traditional methods in predicting ad success? The current study is an attempt to address this question. Our objectives are twofold: First, we explore how measures from commonly used neurophysiological methods tap into higher-level constructs commonly used in advertising research (attention, affect, memory, and desirability). Specifically, we aim to study the relationships among neurophysiological measures as well as their relationship to traditional measures. Second, we aim to explain the variance in real-life advertising success (captured using market response models) using the various neurophysiological measures relative to traditional advertising measures.

The outline of this article is as follows. We first provide a brief literature review on advertising research and introduce the four key constructs in advertising research. Second, we provide an overview of the proposed neurophysiological methods and measures and their relationship to the four advertising constructs. Third, we describe the research design and experimental protocol and discuss the empirical relationships among all our measures. We then describe how we estimated the elasticities for the 30-second television ads, followed by the empirical analysis linking all measures to ad elasticities. Finally, we discuss the study's contributions and implications for researchers and practitioners.

BACKGROUND, LITERATURE REVIEW, AND THEORY DEVELOPMENT

We review the key constructs examined in advertising research by discussing how they have been assessed using traditional self-reported measures. We then introduce the proposed neurophysiological methods and discuss how they can capture traditional advertising constructs more directly and objectively.

Constructs in Advertising Research

Advertising research has long relied on self-reported measures (Biel and Bridgwater 1990; Du Plessis 1994; Haley and Baldinger 2000; Poels and Dewitte 2006; Smit, Van Meurs, and Neijens 2006; Walker and Dubitsky 1994). Researchers and practitioners have used traditional methods of copy testing, such as focus groups and surveys, to collect responses toward ads. These measures are inexpensive, accessible, quick, and relatively simple to analyze; in addition, they offer insight into consumer brand attitudes and preferences as well as into how different ad executions moderate these responses. The common traditional measures can be broadly classified into measures focused on (1) ad execution (e.g., liking, excitability, recall) and (2) the product featured in the ad (e.g., attitudes, purchase intent), and both have been used to explain advertising success.

The early AIDA (attention, interest, desire, action) model argues that every advertising process begins with capturing attention, followed by information assimilation and compre-

[1]The neuromarketing–advertising practice link received a boost through the Advertising Research Foundation's (ARF's) "Neuro 1.0" project (Stipp and Woodard 2011). The goal of the project was to "help the (advertising) industry learn how best to apply the capabilities of neuromarketing to real marketing issues and decisions" (Stipp and Woodard 2011, p. 5).

[2]The impetus for this study came from the follow-up project to Neuro 1.0. The project, called "Neuro 2.0," was initiated by ARF and underwritten by six corporate sponsors that provided both ads for testing and time-series data for estimating elasticities.

hension, which leads to desirability, followed by action (Strong 1925). This and other "hierarchy of effects" models have been the backbone of advertising research for the past few decades (Barry and Howard 1990). Recent research has extended this notion of hierarchy (temporal sequence) by shifting focus to a set of core constructs: attention, affect, memory, and desirability. These core constructs can affect advertising success independently or in combination (Haley and Baldinger 2000; Morwitz, Steckel, and Gupta 2007; Walker and Dubitsky 1994). We begin with a review of these four constructs and discuss how they are assessed using traditional measures.

Attention. Attention is defined as the ability to focus on certain aspects of the environment while ignoring others. Advertising researchers often refer to attention as the ability to attract focus to an ad. Common measures of attention are liking, informativeness, excitability, and relevancy of the ad (Biel and Bridgwater 1990; Brown and Stayman 1992; Schlinger 1979; Smit, Van Meurs, and Neijens 2006). Newer methods, such as eye tracking, provide more direct measures of attention, as we discuss subsequently. Specifically, researchers make a distinction between endogenous ("top-down") attention, in which specific aspects of the ad are explicitly selected and processed, and exogenous ("bottom-up") attention, in which features of the stimulus attract attention and processing. We argue that such distinctions, while critical, cannot be captured using traditional self-reported measures.

Affect. Emotion refers to a relatively brief episode of coordinated brain, physiological, and behavioral changes that facilitate a response to an external or internal event of significance (Davidson, Scherer, and Goldsmith 2009). Affect, though often used as a synonym for emotion, refers to the outward expression of an emotion. We contend that affect, in the context of advertising, can be broadly classified into two dimensions: valence (relative pleasantness/unpleasantness) and arousal (physiological and subjective intensity). In the AIDA model, emotions and affect were considered merely a means for attracting attention. Thus, they have often been inferred with self-reported measures such as liking and excitability (Poels and Dewitte 2006; Walker and Dubitsky 1994). These measures represent post hoc introspection about affect experienced from an earlier stimulus and thus could be distorted by a variety of factors, including higher cognitive processes. Neurophysiological methods, in contrast, provide a more direct measure of affect, as we discuss subsequently.

Memory. Memory refers to the mechanisms by which past experiences influence behavior. Therefore, memory is often associated with encoding (which occurs during the past event), consolidation (which occurs during the intervening period), and retrieval (which occurs at a future time). Retrieval success is often used as a proxy for the depth to which information was encoded (Mandler 1980). Advertising research, like most memory research, has focused on the retrieval aspects to evaluate the quality of ads. The emphasis has been on two retrieval measures in particular: recall, in which participants generate the target with partial or no cues, and recognition, in which participants distinguish the targets from novel distractors (Du Plessis 1994; Singh, Rothschild, and Churchill 1988). Although better memory is often attributed to better ad processing, these measures do not necessarily distinguish between processing due to encoding (when the ad is presented) and retrieval (when the recognition test is performed).

Desirability. In traditional advertising research, desirability refers to the extent to which people desire the product featured in the ad. Marketing managers routinely use measures such as purchase intent as a strong correlate of desirability and subsequent market behavior. To quantify the specific effects of an ad, researchers measure purchase intent as a change in the level of desirability for the product pre- and post-exposure to the ad. This shift measure for purchase intent is often weakened by the varying amounts of brand equity. Because popular brands tend to have higher premeasure scores, it is important to account for this bias before making any judgments about shifts in desirability. Walker and Dubitsky (1994) propose a method by which change scores are normalized by using a baseline predicted average result (PAR) score to remove brand-specific effects not associated with ad exposure. Yet broader concerns still remain about the relationship between these intent measures and subsequent purchasing (Morwitz, Steckel, and Gupta 2007). Consumers are not capable of perfectly predicting the future, either in terms of how they represent their intentions or how these intentions will change over time. The strength of the predictability also depends on the context, novelty, and specificity of the products concerned (Morwitz and Fitzsimons 2004). Therefore, we contend that the extent of reward-related activation in the brain during the actual ad provides a better and more direct measure of desirability, as we describe next.

Newer Methods in Advertising Research

In the past decade, there has been a burgeoning use of neurophysiological methods to understand consumer behavior. In this section, we provide a brief introduction to the methods used in this study. Other sources are available for more detailed reviews of each method (e.g., Dimoka 2012; Huettel, Song, and McCarthy 2008; Potter and Bolls 2012; Shaw 2003; Wedel and Pieters 2008).

Implicit measures. Despite their popularity, self-reported measures are inherently subjective and incomplete because they only capture conscious, declared opinions (Micu and Plummer 2010). As a result, implicit testing has emerged as an alternative to capture the unconscious nature of consumer preferences, attitudes, and information processing (Greenwald, McGhee, and Schwartz 1998). The Implicit Association Task (IAT) is a commonly used measure that captures the strength of association among concepts and avoids tapping into the consumer's conscious thought. Specifically, differences in response latencies for brands paired with positive and negative words in IAT have been used as a measure of emotional valence (Dimofte 2010).

Eye tracking. Next to traditional and implicit measures, eye tracking is perhaps the most accessible method for capturing ad response. Eye tracking has a high temporal resolution (60–120 Hz) and provides insight into temporal processes. Compared with old camera-based systems (with chin rest and head straps), modern eye trackers use an optical camera to identify the position of the pupil and cornea using infrared/near-infrared light that evokes corneal reflection. By tracking participants' gaze when viewing ads, we can capture not only which information was processed but also the order and duration of these processes. Eye tracking has been used as a direct measure of attention. For example, bottom-up factors, such as color and luminance, have a strong effect on initial eye movements (Leven 1991). In addition, the percentage of valid fixations (total amount of time eyes are focused on the ad) provides an index of overall attention or engagement with the ad (relative to distractions). The number of fixations and mean dwell times provide a measure of the depth to which information within an ad is processed (Venkatraman, Payne, and Huettel 2014). Longer dwell times and fewer fixations represent more detailed processing (Horstmann, Ahlgrimm, and Glockner 2009). Finally, eye tracking can also measure pupil dilation[3] (physiological response of the sympathetic nervous system), which provides additional insight into the degree of arousal following an external stimulus (Hess and Polt 1960).

Biometrics. Biometrics refers to the physiological or automatic responses to an external stimulus. Biometrics has become increasingly popular in marketing and advertising research because they can provide insight into unconscious processes and affect (Potter and Bolls 2012). Common physiological responses include heart rate, breathing, and skin conductance.

Heart rate, also called pulse, is the speed of the heartbeat and is typically measured with an electrocardiogram, which measures the electrical activity of the heart using external skin electrodes. Heart rate is controlled by two antagonistic systems: the sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS) (Potter and Bolls 2012). The SNS (termed the "fight-or-flight system") represents the body's automatic response to external stimulus. Activation of this system increases heart rate, also called heart rate acceleration, which provides an independent measure of arousal (Wang, Lang, and Busemeyer 2011). Conversely, the PNS (termed the "rest-and-digest system") refers to a calm and relaxed state that is characterized by slower heart rate, or heart rate deceleration. Increased heart rate deceleration in response to an ad implies increased ability to focus on the ad and thus provides an independent measure of attention (Lang et al. 1999).

Breathing frequency, or respiration rate, refers to the number of breaths taken within a fixed amount of time, typically 60 seconds, to yield a breaths-per-minute (BPM) measure. Activation of the SNS leads to an increase in respiration rate, which can then be used as a measure of arousal. Breathing also influences the SNS/PNS by temporarily blocking PNS influence on heart rate, resulting in increased heart rate, but subsequent exhalation removes this block and decreases heart rate. The undulation in heart rate caused by respiration is called respiratory sinus arrhythmia, which has also been used as a measure of arousal and affective processes (Potter and Bolls 2012).

Skin conductance response (SCR), also known as electrodermal response, occurs when the skin transiently becomes a better electrical conductor due to increased activity of the eccrine (sweat) glands following exposure to certain stimuli (Potter and Bolls 2012). Skin conductance is frequently used as a tool to measure tonic activity of the SNS. Due to the nature of physiological responses, SCR is also preceded by a small latency (delay). Skin conductance amplitude and response latency provide direct measures of arousal when watching an ad, unlike self-reported measures, which are often based on introspection at a later time. Still, SCR cannot reliably indicate emotional valence (Potter and Bolls 2012).

EEG. Perhaps the most commonly used neuroscience method in advertising research (Wang and Minor 2008), EEG can reveal variations in electrical signals of cortical brain regions as a function of internal or external variables. These variations are recorded at different frequencies—delta rhythms (<4 Hz), theta rhythms (4–7 Hz), alpha rhythms (8–12 Hz), and beta rhythms (15–30 Hz)—and correspond to different physiological phenomena. Electroencephalography provides high temporal resolution but low spatial resolution because it is restricted to measuring only cortical brain activity. Here, we focus primarily on the alpha frequency band, which is inherently inhibitory and thus inversely related to underlying brain activity (Jensen and Mazaheri 2010; Shaw 2003). Specifically, we focus on two measures: occipital alpha activity and frontal asymmetry. The occipital alpha measures the extent of activation/gating in the visual system and thus provides an index of visual processing and exogenous attention (Foxe and Snyder 2011; Jensen and Mazaheri 2010). We predicted that the more effective ads would have reduced occipital alpha. Similarly, the relationship between affect and hemispheric asymmetries in frontal brain activity has a long history in psychology and neuroscience (Davidson 2004; Demaree et al. 2005; Harmon-Jones, Gable, and Peterson 2010). This frontal asymmetry measure (ln[F4] – ln[F3]) argues for greater responses in the alpha band frequencies for positive stimuli in the left hemisphere (F3) and negative stimuli in the right hemisphere (Davidson et al. 1990; Tomarken, Davidson, and Henriques 1990). However, others have argued that greater activation in the left hemisphere (smaller alpha) merely reflects approach motivation, independent of emotional valence (Harmon-Jones et al. 2006; Sutton and Davidson 1997). In this study, we expect that the more effective ads will be associated with higher values of frontal asymmetry (approach behavior).

fMRI. Functional magnetic resonance imaging is a noninvasive method that localizes and tracks changes in blood oxygenation during cognitive tasks (Ogawa et al. 1990). The blood oxygenation level–dependent contrast is based on the fact that hemoglobin has different magnetic properties depending on its oxygenation state. Because neural activity following a specific task utilizes oxygen within specific areas of the brain, the brain vasculature responds by increasing the flow of oxygen-rich blood into the region. This leads to a localized increase in blood oxygenation level–dependent signal intensity in that region of the brain, which is then measured using high-field magnetic resonance scanners (Huettel, Song, and McCarthy 2008). Accordingly, fMRI provides an indirect and correlative measure of local brain activity at high spatial resolution (approximately 1

[3]We do not analyze pupil dilation further in this study, because the ads were not controlled for luminance and brightness, which are known to affect pupil dilation.

mm^3) and good temporal resolution (approximately 2–5 seconds).

Neural activations can be used as a direct measure of exogenous and endogenous attention. Exogenous attention is measured through activation in the primary visual cortex (greater visual processing) and amygdala (arousal). In contrast, top-down attention depends on goals, internal states, and expectations (e.g., health cues could help modulate choices of different food items) and is associated with activation in the dorsolateral prefrontal cortex (dlPFC). The dlPFC is the "executive" part of the brain that helps process contextual information (Hare, Malmaud, and Rangel 2011; Miller and Cohen 2001).

The amygdala has been a focus of research on affect and emotions because it is a key part of the limbic system and connects to subcortical structures that process autonomic functions (Pessoa and Adolphs 2010; Phelps 2004). Across several studies, the amygdala has consistently been shown to be involved in various aspects of emotional processing. Specifically, the magnitude of its activation is often related to affective intensity and has been described as being greater for negative than for positive stimuli (Critchley et al. 2005; Dimoka 2010; Sabatinelli et al. 2005). However, lesion studies have indicated that the amygdala may play a more important role in emotional arousal than in valence (Glascher and Adolphs 2003).

It is advantageous to use fMRI to measure memory because it provides a direct measure of the strength of encoding during the ad. For example, we can explicitly identify brain regions that show greater activation for stimuli that were remembered versus forgotten. The hippocampus has been shown to be critical for memory across many neuroimaging studies (Zola-Morgan and Squire 1993), and lesions to this region affect a person's ability to form new memories and associations (Corkin 1984).

Activation in the ventromedial prefrontal cortex (vmPFC) and ventral striatum are viewed as key measures of desirability. The vmPFC has consistently been linked to willingness to pay for a wide range of branded products across different studies (Plassmann, O'Doherty, and Rangel 2007; Plassmann, Ramsøy, and Milosavljevic 2012). The ventral striatum is the primary dopaminergic target in the brain and thus plays an important role in the prediction and consumption of rewards (Knutson, Delgado, and Phillips 2010; Levy et al. 2011). The ventral striatum also plays an important role in wanting, which refers to motivation or approach behavior toward rewards. Although wanting is often correlated with liking (hedonic value of reward), the two concepts can be distinguished through the manipulation of dopamine levels (Berridge 2007). Recent research has shown that activation in the ventral striatum during product evaluation is the strongest predictor of subsequent purchases (e.g., Berns and Moore 2012; Knutson et al. 2007).

We summarize the various neurophysiological measures under the four proposed constructs (attention, affect, memory, and desirability) in Table 1. A more detailed version with references is available in Web Appendix A. As Table 1 shows, these constructs can be assessed using measures from different methodologies. The approach taken in this study is first to compare across these measures along the four key advertising constructs and then to determine which of these neurophysiological measures can predict advertising success.

Table 1

METHODS, MEASURES, AND CONSTRUCTS

	Constructs			
Measures	*Attention*	*Affect*	*Memory*	*Desirability*
Traditional				
Liking	x			x
Excitability	x	x		
Familiarity			x	
Recognition			x	
Purchase intent				x
IAT				
IAT valence		x		
IAT memory			x	
Eye Tracking				
Fixation count	x			
Dwell time	x			
Pupil size		x		
Biometrics				
Heart rate deceleration	x			
Respiratory sinus arrhythmia	x			
Heart rate acceleration		x		
Skin conductance		x		
EEG				
Occipital alpha	x			
Frontal asymmetry		x		
fMRI				
dlPFC	x			
vmPFC	x			x
Amygdala		x		
Hippocampus			x	
Ventral striatum				x

EXPERIMENTAL METHODS

All studies were approved by Temple University's Institutional Review Board. We tested a total of 37 ads in the main study. All ads were 30-second television ads drawn from six companies (described in detail subsequently) and included 15 unique brands. Participants were recruited from a large city in the U.S. Northeast through online and print ads. Interested participants were required to fill out an online prescreening questionnaire at least two days before participating in the study. In addition to basic demographics (gender, ethnicity, age, employment status, and income), we also collected information about television and television-ad-watching habits during the prescreening. Participants who did not watch television or television ads were excluded from the study. To measure the participants' predisposition to products and brands, we showed them images of brands featured in the study and collected information about their product familiarity, purchase intent, usage intent, and recommendation intent. To minimize biases, we included other products from competitors as part of the prescreening questionnaire. For the main study, we collected data from a total of 277 participants across four separate phases. The experimental protocol was largely identical across phases, except for minor methodology-specific modifications. Next, we describe each of the four phases briefly. Additional details are available in Web Appendix B.

Phase 1: Traditional and Implicit Measures

A total of 186 participants (86 women; mean age = 39 ± 14 years) completed Phase 1. All studies were conducted in a laboratory for greater experimental control. A lab assistant briefed participants and obtained their signed informed consent before any data collection. Participants were then seated in front of a computer with headphones. All stimuli were presented using E-Prime 2.0 (Psychology Software Tools), and responses were captured using mouse and keyboard. Participants were provided $15 as compensation for this phase of the study.

We summarize the basic protocol in Figure 1, Panels A and B.[4] Because some brands were repeated across ads, we split our protocol into two pods. In pod 1, participants watched the ads from each of the 15 unique brands. After a five-minute anagram distractor task, they retrieved as many brands as possible from the ads they had just watched, as a free recall measure. The participants then completed pod 2, consisting of the remaining 22 ads. Ads were rotated within each pod across participants. After each ad, participants were asked a series of ten self-report questions, which were drawn primarily from "The ARF Copy Research Validity Project" (Haley and Baldinger 2000) and the "Advertising Research Foundation Copy Research Workshops" from the early 1990s. They included five measures of the ad (liking, excitability, relevance, informativeness, and familiarity) and four measures of products featured in the ad (purchase intent, recommendation intent, usage intent, and familiarity). All questions are listed in Web Appendix B. To assess desirability, we measured the change in product familiarity, purchase intent, usage intent, and recommendation intent from the baseline measures obtained during prescreening. We normalized these change scores using a PAR measure to remove effects that were brand specific and not associated with ad exposure (Walker and Dubitsky 1994).

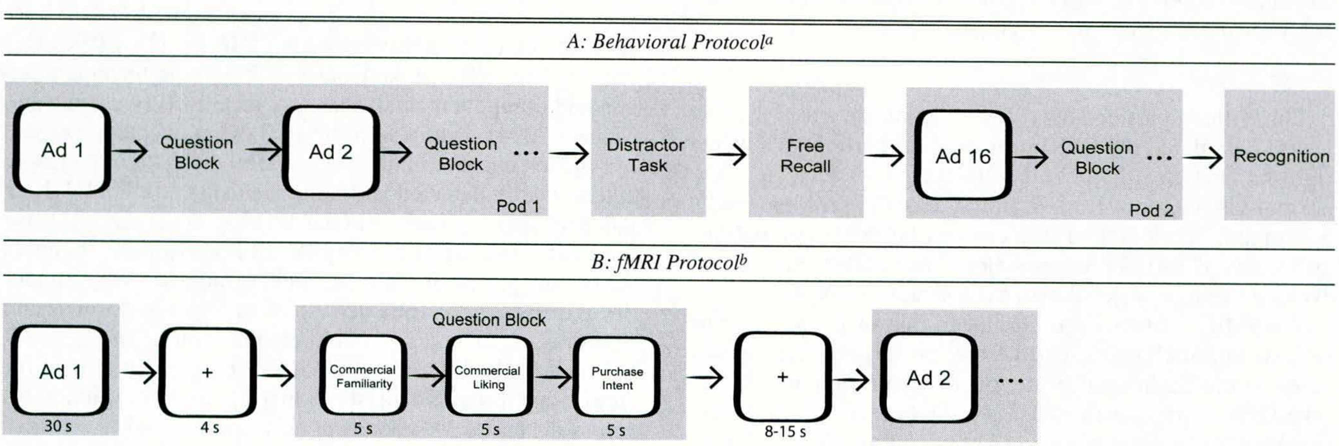
After viewing all ads, participants were given a five-minute break before a surprise recognition test. We identified two salient moments for each ad based on internal pretesting. We then used one of the screenshots (unbranded moment) interspersed with foils (screenshots drawn from similar products). Participants were asked to indicate whether each of the screenshots was old (from ads they had seen in the session) or new on a six-point scale that included confidence measures (Web Appendix B). To calculate the hit rate, we converted the responses into a simple binary scale.

For 80 (41 women) of the 186 participants, we administered a modified version of the IAT (Greenwald, McGhee, and Schwartz 1998) after the recognition test.[5] Similar to the original IAT, participants were asked to sort stimuli into different categories as quickly as possible. Participants categorized words as either positive (e.g., love) or negative (e.g., death) and categorized images as representing indoor or outdoor scenes (Web Appendix B). The images were salient, unbranded screenshots drawn from the ads, interspersed with foils selected from competitor ads. We used the difference in response latencies to ad images versus foil images as an implicit measure of memory (previously seen images are likely to be retrieved more quickly). We refer to this measure as IAT_Memory. The difference in response

[4]We carried out two pilot studies to address concerns that (1) answering self-reported measures immediately after ad could bias memory measures and (2) the large number of ads tested could affect participants' engagement and recognition scores (for details, see Web Appendix B).

[5]Participants also completed a second implicit task called the Affect Misattribution Procedure (Murphy and Zajonc 1993) after the IAT. However, because the Affect Misattribution Procedure focused only on the constricted set of unique brands, we do not discuss it further.

Figure 1
EXPERIMENTAL PROTOCOL



[a]Participants viewed 37 30-second television spots, separated into two pods. In Pod 1 (15 ads), all ads represented unique brands. In Pod 2, the remaining ads were presented. Ads were rotated within each pod across participants. A question block followed each ad, in which a series of self-reported measures was obtained. A recall task was administered for the unique brands in between the two pods, following a distractor task. Surprise recognition tests were administered at the end of the study.

[b]For the fMRI protocol, the two pods were divided into five runs. Each run was exactly eight minutes long. Each 30-second ad was followed by a 4-second fixation (a cross presented in the center of the screen) and sequential presentation of three questions. Participants had up to 5 seconds to answer each question. If a response was recorded before the 5-second limit, the remaining time was filled with a fixation. There was a variable intertrial interval of 8 to 12 seconds before the next ad. Eight ads were presented in each run.

latency between when each image was paired with a positive or negative word served as an implicit measure of emotional valence toward each ad. We refer to this measure as IAT_Valence. We excluded 22 participants who had error rates greater than two standard deviations from the mean (Greenwald, McGhee, and Schwartz 1998), resulting in a total of 58 participants for subsequent analysis. The excluded participants had greater difficulty in reversing responses between blocks, possibly because of fatigue from the lengthy experimental protocol.

Phase 2: Eye Tracking and Biometrics

A total of 29 participants (11 women; mean age = 33 ± 10 years) completed the eye-tracking and biometric studies. After a briefing similar to the traditional phase, participants sat in front of a Tobii T60XL eye tracker and were affixed with BIOPAC (MP150) BioNomadix wireless physiology devices for collecting skin conductance, heart rate, and breathing data. Stimuli were presented using E-prime 2.0. The protocol was similar to Phase 1, with additional breaks to mitigate participant fatigue.

For eye tracking, information about fixations and gaze locations were exported from Tobii and analyzed using in-house scripts in MATLAB. Biometric data were preprocessed using the Acqknowledge 4.0 package. For heart rate data, raw tonic data were analyzed using Acqknowledge's Heart Rate Variability procedure. Although typical heart rate analysis focuses on the low-frequency (.04–.15 Hz) and high-frequency (.15–.40 Hz) components as measures of accelerations and decelerations, they often require events of longer durations for reliable estimations compared with the 30-second ads used here. Therefore, we created a coding system to identify heart rate accelerations and decelerations using the phasic heart rate signal. We coded any positive shifts from the baseline (measured before the start of each ad) as heart rate acceleration and any negative shifts from baseline as heart rate deceleration. We analyzed event-related SCR data using Acqknowledge's built-in exploratory data analysis. Additional details about the setup and specific analyses are available in Web Appendix B.

Phase 3: fMRI

Thirty-three participants (15 women, mean age = 29 ± 8 years) completed the fMRI protocol. All participants were right-handed, healthy people with normal or corrected-to-normal vision and were free of any hearing problems. All participants provided written consent before participating and received $40 as compensation. Four participants were excluded from analysis due to excessive movement.

The fMRI protocol was similar to that of Phase 1, with several minor changes. Unlike the ten self-reported measures in the traditional protocol, fMRI participants were asked only three questions to keep the overall duration reasonable. We selected measures about ad familiarity, liking, and purchase intent because these measures showed the most variability in our preliminary analysis. We then divided the 37 ads into five runs with breaks between each run to reduce fatigue. To keep run lengths consistent (eight videos per run), we added one filler ad that was always the last video in the second run as well as two others at the end of the fifth run. After the first two runs, which consisted of ads with unique brands, participants rested for four minutes with eyes open and fixated on a cross in the center of the screen. This rest period acted as a distractor for the subsequent free recall test that was administered in the scanner through an intercom. We summarize the timing for the fMRI protocol in Figure 1, Panels A and B, and provide details about the fMRI sequence in Web Appendix B. After the fifth run, a recognition test was administered outside the scanner on a laptop.

We constructed a first-level general linear model (Friston et al. 1994) for each of the ads for each participant. This model consisted of one regressor (30 seconds) for each of the ads. All three traditional measures (familiarity, liking, and purchase intent) were collapsed across ads into three regressors per run. Motion parameters were included in both models as an effect of noninterest. We then constructed a second-level model for each of the 37 ads as one-sample t-tests in SPM8 using contrast images from the first-level model for each of the 29 participants (Berns and Moore 2012). Then, we built a third-level model (also a one-sample t-test) using contrast images from the second level and additional covariates (from within- or out-of-sample traditional measures). For analysis with covariates, statistical images had a threshold of $p < .001$, uncorrected. We also preselected an independent set of four regions on the basis of their established role in measuring the core constructs, for a region-of-interest (ROI) analysis (Web Appendix B). For each ROI, the parameter estimates for each ad were obtained from the second level using MarsBaR toolbox for SPM (Brett et al. 2002). Unless specified otherwise, all brain activations in this study refer to these preselected ROIs and are not from any specific models.

Phase 4: EEG

We obtained high-density EEG data from 29 participants (15 women; mean age = 25 ± 6 years). We used a 129-channel HydroCel Geodesic Sensor Net (Electrical Geodesics Inc.) with a Cz reference to record EEG data. The protocol was almost identical to the fMRI protocol (the liking measure was not obtained due to a coding error). Raw EEG data were first filtered using a bandpass filter (HP:.01 Hz; LP: 40 Hz) and rereferenced to linked mastoids before performing independent component analysis using EEGLAB (Delorme and Makeig 2004). Artifacts (horizontal eye movement, vertical eye movement, eye blinks, and general discontinuities) were automatically detected and removed using ADJUST 1.1, an independent plug-in for EEGLAB (Mognon et al. 2011). We then extracted alpha activity (8–12 Hz) from 17 channels for the final analysis (for additional details, see Web Appendix B). Extracted alpha activity was log-transformed and baseline corrected for each channel on a moment-to-moment basis. We then estimated the aggregate mean for frontal asymmetry (ln[F4] – ln[F3]) and occipital (Oz) across the entire 30 seconds of each commercial.

RELATIONSHIP AMONG MEASURES OF ADVERTISING EFFECTIVENESS

Relationship Among Traditional Measures

We first examined the relationship among the traditional self-reported measures. We restricted this analysis to the 186 participants from Phase 1. Web Appendix C summarizes the pairwise rank correlations across ads among the different measures. We found significant positive correlations among the various ad-related measures such as liking, familiarity, relevance, and informativeness. We also found significant correlations among the various product-related measures—namely, changes in purchase intent, usage intent, recommendation intent, and familiarity with the products featured in the ad. The ad-related measures were also correlated with the product-related measures. Finally, recognition was significantly correlated with excitability and liking.

We next categorized the traditional advertising measures using factor analysis. Using a Varimax rotation, we found that the 11 measures loaded mainly onto three factors: one factor loaded strongly on all the 6 ad-related measures; the second factor loaded on all product-related change measures and weakly on liking, excitability, and relevance; and the third factor loaded strongly on recognition and weakly on excitability and liking (Web Appendix C). Therefore, we selected liking, change in purchase intent, and recognition as key traditional measures for further analysis because they loaded highly onto one of the three different factors and were most consistent with prior copy testing research.

Relationship Among Traditional Measures Across Samples

A different set of participants completed each of the four experimental phases, so we analyzed the consistency in the traditional measures across phases. The four measures collected across all phases were ad familiarity, liking, purchase intent, and recognition. For each of these measures, we calculated a measure of internal consistency (Cronbach's alpha) across phases. We found strong consistency for liking (α = .916), familiarity (α = .796), change in purchase intent (α = .739), and recognition (α = .931). This suggests that the self-reported measures across the various sets of participants were consistent in this study.

Relationship Between Biometric and Traditional Measures

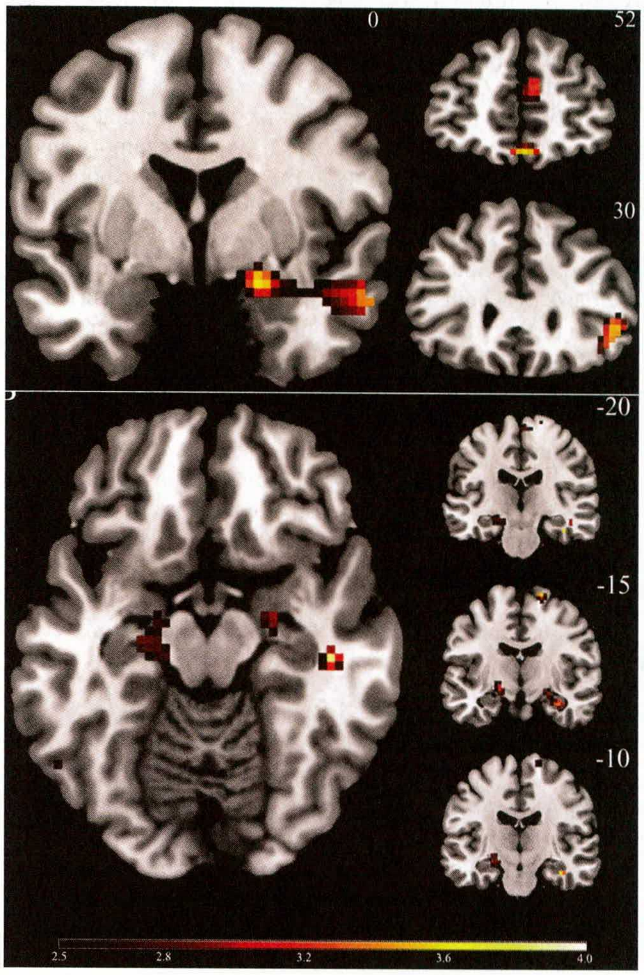
We then examined the relationship between the biometric and traditional measures for all participants in Phase 2 (Web Appendix C). We found that deceleration correlated with liking ($r = .37, p < .05$), recognition ($r = .34, p < .05$), and change in purchase intent ($r = .46, p < .01$). These findings are consistent with deceleration providing an independent measure of increased attention. There was also a negative correlation between heart rate acceleration and deceleration ($r = -.52, p < .001$). Therefore, we focused on heart rate deceleration for the remaining analyses. We did not find any significant correlations between SCRs and any traditional measures. Finally, among the various eye-tracking variables, we found that the percentage of valid fixations was significantly correlated with liking ($r = .38, p < .05$). This is again consistent with the finding that ads that were liked were associated with increased attention and processing. All other associations were not statistically significant.

Relationship Between fMRI and Traditional Measures

We aimed to elucidate the neural correlates of the three key traditional measures: liking, purchase intent, and recognition. Unlike typical fMRI analyses that use aggregate measures from the sample of fMRI participants as covariates, we used the average liking and purchase intent measures for each of the ads across participants from all four phases as covariates to identify regions in the brain that tracked these measures.

Using the liking measure as a covariate, we found significant activations in the right amygdala, dlPFC, and vmPFC (Figure 2, top). Historically, liking has been argued to represent both cognitive and affective processes. The pattern of activations found here is consistent with the presence of these two components: the amygdala represents affective processing (Pessoa and Adolphs 2010; Phelps 2004), and the dlPFC represents cognitive processing (Miller and Cohen 2001). To better understand the nature of this interaction, we used a bootstrapping mediation analysis (Preacher and Hayes 2004) to investigate whether the amygdala activation mediated the effect of liking on the

Figure 2
NEURAL CORRELATES OF LIKING AND MEMORY FOR TELEVISION ADS



Notes: The top panel shows brain regions that positively correlated with an out-of-sample measure of average likability across ads. Using a threshold of $p < .001$, we found significant activation in the right amygdala, right dlPFC, and right vmPFC. In the bottom panel, using subsequent-memory analysis, we found significant activation in the bilateral hippocampus when we compared activity for recognized versus unrecognized ads.

dlPFC. Liking had a significant effect on dlPFC activation ($\beta = .30, p < .01$). However, when introducing amygdala activation as a mediator, the direct effect of liking on the dlPFC became insignificant ($\beta = .07, p = .43$), while the indirect effect through the amygdala was significant, implying that the amygdala fully mediated the effect of liking on the dlPFC. Therefore, we contend that liking leads to increased arousal and affect, which in turn translates to greater top-down attention and cognitive processing. Finally, activation in vmPFC is consistent with Berns and Moore (2012), who find activation in vmPFC to covary with likeability ratings of audio songs.

We next used purchase intent as a covariate to identify brain regions that tracked desirability for the products featured in the ad. We used only the postmeasures of purchase intent ("How likely are you to purchase the product in the ad you just watched?") for this analysis because they are closest to the fMRI activations. We found activation in the vmPFC and more rostral region of the anterior cingulate cortex. Again, activation in the vmPFC is consistent with other studies that have postulated an important role for this region in estimating willingness to pay for products and for product valuation (e.g., McClure et al. 2004; Plassmann, O'Doherty, and Rangel 2007; Plassmann, Ramsøy, and Milosavljevic 2012).

Finally, we identified regions in the brain that tracked recognition. We ran a traditional subsequent-memory analysis focused on the moments in the ad that corresponded to the images used in the recognition test. In a separate model, we included two additional regressors that were each two seconds long and classified as "remember" or "forgot" on the basis of whether the participants correctly classified the image in the subsequent recognition test. We then used a paired t-test at the second level to look for differences in the brain between these two regressors. Consistent with our predictions, we found significant activation in the bilateral hippocampus (Figure 2, bottom). In other words, ads that had higher activations in the hippocampus (stronger encoding) during the initial presentation were more likely to be remembered in a surprise memory test later. These findings are consistent with prior studies that have also shown a strong link between memory-related activation in the hippocampus and brand preferences (e.g., McClure et al. 2004).

Summary of Relationships Across Traditional and Neurophysiological Measures

First, we confirmed the consistency of traditional measures across all four phases. Second, we demonstrated high reliability in the self-reported measures between the various samples used in the study, which enabled us to compare and integrate data across the different methodologies. We then demonstrated relationships of the neurophysiological measures to the appropriate construct (attention, arousal, memory, and desirability) and to one another. For example, we found a strong positive correlation between the SCR amplitude and frontal asymmetry measure ($r = .39, p < .05$) obtained from EEG, suggesting that ads with higher arousal levels as measured by SCR amplitude were also associated with higher frontal asymmetry (greater approach behavior).

Next, we explore how the neurophysiological measures predicted ad effectiveness, measured with market response models. We had market response data for 26 of the 37 ads tested in the study (details highlighted in the following section). Therefore, we restricted the number of measures used in the prediction models a priori. We selected 17 measures[6] on the basis of empirical findings, relationships to the core constructs, and relevance to prior advertising literature. They included four traditional measures (liking, product familiarity, change in purchase intent, and recognition),[7] two implicit measures (IAT_Valence and IAT_Memory), two eye-tracking measures (percentage of valid fixations and total number of fixations), three biometric measures (heart rate deceleration, SCR amplitude, and BPM), two EEG measures (frontal asymmetry and occipital alpha), and four fMRI areas (vmPFC, dlPFC, amygdala, and ventral striatum). Table 2 summarizes the measures, their means across ads, and correlations between them for these ads.

ADVERTISING ELASTICITY ANALYSIS

In the second stage of the study, we aim to link the 17 measures across traditional and neurophysiological methods to market-level response to advertising. Our study addresses the following primary question: Which of the measures explains the most variance in market response to advertising beyond the traditional measures that have been used in theory and practice for many years? To answer this question, we developed a two-step process. In Step 1, we estimate a sales response model by specifying and estimating the market responses to the television ads on a company-by-company basis. In Step 2, given the response parameters estimated in Step 1, we regress these parameters on different subsets of the variables aggregated over participants from the six sets of measures.

Step 1: Estimating Advertising Elasticities

We acquired sales and gross rating points[8] (GRPs) data from four of the seven companies in the study as well as elasticities estimated directly by one of the other companies. We were not able to obtain demand or elasticity data from two of the companies. Given differences in industry types and data availability, we sought a measure of response that would be comparable across the different companies and product categories. "Advertising elasticity" is the percentage change in sales due to a 1% change in the advertising measure being utilized (e.g., expenditures, GRPs) and has been used extensively in the literature. Advertising elasticities have several attractive features. First, they are dimensionless, so they can be estimated independently of the units of analysis. Second, they can be computed for any dependent variable using suitable variable transformations.

[6]We also considered excitability and activation in the hippocampus but subsequently excluded them because they were very highly correlated ($r > .8$) with liking and amygdala, respectively.

[7]Although recall is a popular measure of memory, we did not analyze it here because it was restricted only to the 15 unique brands.

[8]Gross rating points are a measure of the size of an audience for an advertisement. They measure the reach of an ad in terms of the percentage of the target audience exposed multiplied by the frequency with which the exposure occurs.

Table 2

RELATIONSHIPS BETWEEN MODELED VARIABLES

			Traditional				IAT		Eye Tracking		Biometrics			EEG		fMRI		
	M	*SD*	*Liking*	*Familiarity*	*Δ PI*	*Recognition*	*Memory*	*Valence*	*Number of Fixations*	*Percentage of Fixation*	*SCR Amplitude*	*HR Deceleration*	*BPM*	*Oz*	*Asymmetry*	*Amyg*	*dlPFC*	*vSTR*
Liking	3.69	.22																
Familiarity	3.91	.35	.70**															
Δ PI	−.05	.17	.52**	.22														
Recognition	.83	.14	.19	−.13	.06													
IAT memory	−105.81	106.94	.05	.21	.13	−.03												
IAT valence	33.03	160.90	−.21	−.19	−.10	−.22	−.31											
Number of fixations	75.08	6.67	−.26	−.19	−.26	.38	.16	.06										
Percentage of fixation	.74	.03	.29	.13	.02	.06	.06	−.30	.03									
SCR amplitude	1.48	.70	.12	.02	.25	−.12	.28	−.29	−.30	.12								
HR deceleration	11.06	4.33	.37	.21	.19	.31	.06	−.22	−.14	−.11	−.03							
BPM	17.35	.70	−.04	−.30	−.15	−.22	−.32	.06	−.07	.31	−.27	−.23						
Occipital alpha	−.10	.02	−.08	−.08	.18	−.04	.01	−.08	.18	−.12	.09	−.02	−.10					
Frontal asymmetry	.00	.01	.13	.00	−.04	.30	−.15	.14	−.13	.02	.39*	−.02	−.10	−.08				
Amyg	.44	.18	.44*	.27	.08	.29	−.24	−.03	−.14	−.14	−.21	.39*	−.08	.07	.05			
dlPFC	.98	.16	.34	.06	.38	.04	.38	−.12	−.10	−.05	.06	.10	−.01	.02	−.07	.31		
vSTR	.13	.11	−.03	−.21	.14	.03	.11	.05	-.43*	−.09	−.08	.20	.07	.05	.28	.13	.12	
vmPFC	.28	.19	.28	.57**	−.14	−.19	.24	.11	−.23	−.12	−.20	.27	−.22	−.16	.07	.40*	.26	.21

*$p < .05$.
**$p < .01$.
Notes: Δ = (Post – Pre) – PAR; PI = purchase intent; IAT = Implicit Association Test; SCR = skin conductance rate; HR = hear rate; BPM = breaths per minute; Oz = occipital; Amyg = amygdala; dlPFC = dorsolateral prefrontal cortex; vSTR = ventral striatum; vmPFC = ventromedial prefrontal cortex.

The general model, which is estimated by industry for brand i, takes the following form:

$$(1)\qquad DV_{it} = G_{it}\beta_i + Z_{it}\lambda + X_t\alpha + \varepsilon_{it}, \text{ and}$$

$$(2)\qquad G_{it} = \delta G_{it-1} + GRP_{it},$$

where

GRP_{it} = a vector of GRPs for all ad spots at time t for brand i;
G_{it} = a vector of advertising goodwill stock for all ad spots j at time t for brand i;
β_i = a vector of advertising effectiveness for all ad spots j for brand i;
δ = advertising carryover ($1 - \delta$ is the exponential decay rate);
Z_{it} = industry and/or brand-specific variables including media, price, penetration, and so on;
X_t = time-related control variables (e.g., seasonality) at time t; and
ε_{it} = unobserved shock to sales for brand i at time t.

The parameters of interest are the β_i, which, given sufficient data, enable us to estimate the effectiveness of the specific ads and then calculate the ad elasticities. Goodwill was entered linearly, which has the advantage that it does not matter if some ad exposures are aggregated and some are not. For several companies, we had GRP data for the specific ad that was tested in the lab, whereas the GRPs of other ads were aggregated.

If the true dependent variable of interest at the individual level is consumer utility and if we have data that enable us to calculate market share, we could estimate the model given in Equation 1 using the log-odds ratio as the dependent variable, $\log(S_{it}) - \log(S_{0t})$, where S_{0t} is the market share of the outside option and the interpretation of the estimated coefficients would be the effect of the independent variables on consumer utility (assuming a consumer-level Type I extreme value utility shock over which we integrate). In cases in which we did not have competitor information, the utility interpretation was still possible if the market size and share of the outside option did not vary systematically with our control variables. This is because $\log(S_{it}) - \log(S_{0t}) = \log(q/M) - \log(S_{0t}) = \log(q) - \log(M) - \log(S_{0t})$, and so the $\log(M)$ and $\log(S_{0t})$ terms were absorbed in the regression intercept when using log demand as the dependent variable.

A limitation of the analysis is that some of the executions ran for only a short period of time. It was thus impossible to separate short-term and long-term advertising effects. As a result, we created a cumulative advertising term and, drawing on some auxiliary analyses, only estimated long-term elasticities on the basis of Equation 1, with $\delta = .9$. In the following subsections, we provide a brief description of each company's estimates. A total of five firms—two consumer product firms, one large and one multinational financial services firm, and one large Internet travel services firm—provided data for this study. Because the companies provided different data for the estimation of Equation 1 and represented different product types and market conditions, we estimate five separate demand models with different sets of controls. Table 3 presents all results. For additional company-specific details, see Web Appendix D.

Company A. Company A is a large Internet travel services company. Data included weekly national GRP data and market share data between 2010 and 2012 as well as unaided brand recall for the focal company and its major competitors. For this company, we therefore used equivalent models for two dependent variables: log-odds and recall. We estimated the two-equation system using seemingly unrelated regression analysis. By including the additional information in the recall measure, we could better account for the unobserved shocks. The set of control variables included a week-of-year fourth-order polynomial, year-specific week-of-year second-order polynomials, and brand dummy variables.

Table 3

ESTIMATED AD ELASTICITIES

Company	*Ad*	*Estimate*	*SE*
A	Ad 1	.11*	.07
A	Ad 2	.11**	.05
A	Ad 3	.16***	.06
A	Ad 4	.16***	.05
A	Ad 5	.10	.13
B	Brand 1, Ad 1	–.01	.02
B	Brand 1, Ad 2	.09*	.05
B	Brand 2, Ad 1	.09***	.02
B	Brand 3, Ad 1	.18***	.06
B	Competitor Brand 1, Ad 1	.26***	.07
B	Competitor Brand 2, Ad 1	.09***	.01
B	Competitor Brand 2, Ad 2	.09***	.02
C	Ad 1	–.05	.12
C	Ad 2	–.01	.13
C	Ad 3	.13**	.05
C	Ad 4	.41	.27
C	Ad 5	–.12*	.07
D	Brand 1, Ad 1	.33*	.17
D	Brand 1, Ad 2	.12	.16
D	Brand 2, Ad 1	.17***	.05
D	Brand 2, Ad 2	.11	.07
E	Ad 1	.26	—
E	Ad 2	.54	—
E	Ad 3	.23	—
E	Ad 4	.47	—
E	Ad 5	.39	—

*$p < .10$.
**$p < .05$.
***$p < .01$.

Company B. Company B is a consumer products company. Data included weekly national GRP and market share data between 2010 and 2012 for two related product categories. As with Company A, we had competitor demand and GRP data, so we were able to use the log-odds ratio as the dependent variable. To construct the dependent variable, we used sales divided by a price index as the demand variable. The control variables included the price index, the everyday base price, promotion variables (log of the percentage of sales sold with an ad feature with display, feature without display, display without feature, and promotion only), brand-specific week third-order polynomials, and week dummy variables. Company B also provided information on ads from one of its major competitors, which was useful to fully specify the model.

Company C. Company C is a large financial services company. Because we did not have competitor data, we used the log of the sales data for the dependent variable. As with Company A, we had an additional dependent variable that would also reflect explained unobserved demand shocks—namely, the web channel click-through rate. We used log click-through as a second dependent variable in a seemingly unrelated regression. The control variables included the Dow Jones Industrial Average, a week-of-year fourth-order polynomial, and a week third-order polynomial.

Company D. Like Company B, Company D is a large consumer products company. Data included weekly sales and advertising GRP between 2009 and 2012 for four advertising executions. As with Company B, we used sales divided by a price index as the dependent variable. Control variables included the price index and advertising GRP for untested advertising executions.

Company E. Company E is a multinational financial services company. Unlike the other four companies, this company provided its own advertising elasticity estimates.[9]

Estimates

The estimates of the advertising elasticities for all tested ads, which would be used as the dependent variables in the second-stage regression, appear in Table 3. The only negative elasticity estimates are small and not significant; in these cases, we replaced the estimate with zero. Three of the five ad elasticities from Company A are significant at 10%, as are five of the seven from Company B, one of the five for Company C (due to data limitations), and one of the two for Company D. We do not know the significance of the elasticities provided directly by Company E. The mean of the positive, significant elasticities was .14, which is within the range found by many studies and meta-analyses of advertising effectiveness (e.g., Sethuraman, Tellis, and Briesch 2011).

Step 2: Neurophysiological Predictors of Ad Elasticities

In Step 1, we recover the long-term effectiveness of advertising for brand i and ad j (the β_{ij}). In Step 2, we estimate the effects of the various multimethod measures on the effectiveness of television advertising on sales (i.e., the ad elasticities). Let η_{ij} be the ad elasticity for brand i and ad j. We transformed the elasticities for the second stage to use logs of elasticity because we could control for proportional differences in ad effectiveness across industries' brand dummy variables. In practice, we use $\log(.1 + \eta_{ij})$ to prevent taking the log of zero. The general form of the model utilized is

$$\log(.1 + \eta_{ij}) = W_{ij} + N_{ij} + \gamma_i + \xi_{ij}, \tag{3}$$

where W_{ij} represents the traditional measures for ad j by brand i (including company-specific effects for purchase intent), N_{ij} includes the nontraditional measures (including implicit measures, eye tracking, biometrics, EEG, and fMRI), and γ_i are company fixed effects.

We had to make several accommodations to the data and results to estimate Equation 3 because of the limited degrees of freedom. Importantly, as noted previously, we used the reduced set of variables for each category of measures, and in addition, we created aggregated values of the measures by taking the means across the relevant set of respondents. Finally, we ran the regressions separately using different sets of measures because we did not have a sufficient number of observations to include all measures simultaneously in a single regression. However, we contend that such an analysis is relevant for practitioners, who are highly unlikely to invest in all the methods at once and have all these measures.

Our primary goal was to investigate which of the set of measures best explains the variation in advertising elasticities beyond traditional measures. However, we first ran a set of regressions with each set of variables (traditional, IAT, biometrics, fMRI, and EEG) separately with individual-company dummy variables to control for fixed-effect differences among companies. Although not a focus of our analysis, these results (Web Appendix E) show that the traditional variables were by far the best predictors of ad elasticities. They produced a 72% improvement in adjusted R^2, beyond the company dummies.

To assess which measures best improved the explanation of the advertising elasticities beyond these traditional measures, we included each set of nontraditional measures with the traditional measures and company dummies in separate regressions. We then conducted an F-test and assessed whether each method adds a significant explanatory power after controlling for the traditional measures. We present the results in Table 4. We found that when we controlled for traditional measures, only fMRI measures were significant predictors of ad elasticities ($p < .011$). Consistent with this result, fMRI measures were the only variables to produce a positive percentage increase in adjusted R^2. Table 4 also presents the parameter estimates of the relationships between the individual measures and the market-level advertising elasticities, controlling for brand heterogeneity and the traditional measures. Notably, the only significant result is the positive impact of the activation in ventral striatum.

These results suggest that for researchers interested in utilizing one physiological approach beyond the traditional self-reported measures, fMRI would be the best candidate. To test whether some of the other measures explain the same variance as traditional measures, we ran additional regressions with each of the sets of measures in isolation, without controlling for traditional measures. We found that eye tracking and EEG measures were moderate predictors of ad elasticities (for details, see Web Appendix E). Therefore, it is likely that eye tracking and EEG measures could potentially explain much of the same variance in ad elasticities as the traditional self-reported measures.

DISCUSSION

In the past decade, a new industry has rapidly grown around neuroscience applied to marketing, as marketing practitioners increasingly look to neuroscience methods to better understand consumer behavior and advertising. Yet healthy skepticism exists in both academia and practice about the contribution and value of these methods to marketing. This is the first study to provide a framework for how academic research on neuroscience can inform advertising practice. Using a unique experimental protocol, we obtained multiple measures of advertising effectiveness across the six most commonly used methods (traditional self-reports, implicit measures, eye tracking, biometrics,

[9]This company did not disclose exact model and method for estimating their advertising elasticities, which might introduce an inconsistency issue. Therefore, as a robustness check, we conducted another set of analyses by dropping observations for this company (Web Appendix E).

Table 4

EFFECTS OF TRADITIONAL AND NEUROPHSYIOLOGICAL MEASURES ON ADVERTISING ELASTICITIES BEYOND TRADITIONAL MEASURES

	Model 1 IAT		*Model 2 Eye Tracking*		*Model 3 EEG*		*Model 4 fMRI*		*Model 5 Biometrics*	
	Est.	*SE*	*Est.*	*SE*	*Est.*	*SE*	*Est.*	*SE*	*Est.*	*SE*
Company Dummies										
Constant	1.90	2.36	4.10	3.38	2.36	2.47	1.74†	1.20	2.31	3.64
Company A	–.708**	.234	–.706**	.257	–.549*	.271	–.851***	.138	–.652†	.370
Company B	–.582**	.222	–.604**	.236	–.418*	.211	–.508***	.099	–.470	.499
Company C	–.697†	.423	–.893*	.437	–.777*	.416	–.933**	.295	–.670	.582
Company D	–.142	.245	–.145	.142	–.297	.230	–.072	.159	–.122	.411
Company E	.161	.180	.024	.229	.107	.198	–.152	.157	.107	.229
Traditional Measures										
Liking	.091	.560	–.363	.572	–.548	.585	–.123	.417	.012	.911
Familiarity	–.286	.345	–.205	.308	.121	.263	–.178	.340	–.148	.302
Recognition	–1.53***	.401	–.925	.720	–1.04*	.506	–1.51***	.359	–1.50†	.833
Company A × PI	2.70**	.842	1.82†	.981	2.16*	.969	3.09***	.822	2.33	1.52
Company B × PI	4.17**	1.74	4.22**	1.78	5.76***	.884	8.29***	1.42	4.99†	3.18
Company B's Competitor × PI	.030	.758	.505	.921	.077	.812	–.022	.571	–.045	1.18
Company C × PI	–.399	.731	–.641	.899	–.844	1.06	–.984†	.655	–.777	1.30
Company D × PI	–1.57†	.986	–1.01*	.508	–.001	.674	–1.79**	.516	–.824	.683
Company E × PI	3.12*	1.58	.374	2.82	.882	1.77	.300	1.51	1.90	2.94
Implicit Measures										
IAT memory	8.01e-4	6.42e-4								
IAT valence	7.68e-5	3.26e-4								
Eye Tracking										
Number of fixations			–.011	.009						
Percentage of fixation			–.724	2.08						
EEG										
Occipital alpha					2.48	1.90				
Frontal asymmetry					3.38	5.09				
fMRI										
Amyg							–.164	.253		
dlPFC							.330	.319		
vSTR							.869**	.239		
vmPFC							.400	.480		
Biometrics										
SCR amplitude									.017	.078
HR deceleration									2.81e-04	.002
BPM									–.052	.121
Adjusted R^2	.580		.498		.521		.856		.378	
Percentage change in adjusted R^2	8.0%		–7.3%		–3.0%		59.4%		–29.6%	
F-test *p*-value	.471		.389		.399		.011		.949	

†$p < .20$.
*$p < .10$.
**$p < .05$.
***$p < .01$.

Notes: The percentage changes in adjusted R^2 and the F-test *p*-values are computed against the model with company dummies and traditional measures. The base category for company dummies is Company B's competitor. IAT = Implicit Association Test; PI = purchase intent; Amyg = amygdala; dlPFC = dorsolateral prefrontal cortex; vSTR = ventral striatum; vmPFC = ventromedial prefrontal cortex; SCR = skin conductance rate; HR = heart rate; BPM = breaths per minute.

EEG, and fMRI). Furthermore, we demonstrated the relative contribution of these measures in predicting advertising elasticities using independent and objective measures of real-world advertising success obtained with marketing-mix modeling. Our findings suggest that neurophysiological methods can explain significantly greater variance in advertising elasticities than traditional advertising methods alone. We discuss the broad implications of these findings next.

Implications for Theory, Research, and Practice

This study makes three important contributions. First, we develop and test a unique multimethodological experimental protocol that examines the same stimuli (television ads) with a variety of traditional and neurophysiological methods that allow for direct comparisons of these methods. Previously, comparisons across methods were inferred from parallel findings across studies using different stimuli and protocols, thereby preventing a direct comparison. This method integration has important implications for both academic research and practice. For academics, this study paves the way for similar efforts in other areas, such as consumer decision making. For practitioners, it provides a proof of concept for the integration of traditional advertising methods with neurophysiological approaches by using a common protocol as well as a novel perspective toward capturing key marketing variables. Integration of evidence from multiple methodologies will also likely lead to the development of better theories and models in marketing that are grounded on biological plausibility, which would ultimately benefit both academics and practitioners.

The second major contribution relates to the examination of interrelationships among the measures obtained from both traditional and neurophysiological methods (biometrics, EEG, and fMRI) because they correspond to the key constructs associated with advertising success (attention, affect, memory, and desirability) (Table 1). In the past, differences in terminology and language may have isolated academics and practitioners. We hope this study (e.g., Table 1) helps clarify some of the pertinent advertising constructs and how they can be measured differently with multiple methods. A first step toward integrating these methods is to demonstrate the commonalities and differences among measures. We show high reliability across samples for the self-reported measures (demonstrating robustness of these measures) and largely consistent patterns of correlations both across the four key constructs (attention, affect, memory, and desirability) and across measures (in support of internal validity). For example, we show strong correlations among liking, number of fixations, heart rate deceleration, and activation in dlPFC, consistent with the higher-level construct of attention. Strikingly, liking was also correlated significantly with excitability and activation in the amygdala, consistent with previous intuition that liking measures both rational cognitive and also affective unconscious components. Our mediation analysis further supports the notion that affective processes may regulate the degree of top-down attention.

It is also important to acknowledge that we did not confirm all expected relationships, possibly because of variability in responses across participants and smaller sample sizes for some of the methods. For example, skin conductance measures did not correlate with any other arousal measure. This could suggest that skin conductance measures may actually capture different aspects of arousal or merely represent a limitation (as we discuss subsequently). We contend, however, that our findings have implications for theory of how various marketing measures relate to one another and what we can learn from such relationships. These findings also provide valuable insight for advertising theory and measurement about the nature of higher-level constructs commonly used in advertising research.

Finally, and perhaps most importantly, this is one of the few studies to demonstrate the relationship between laboratory measures and real-world market outcomes, with obvious implications for practitioners. After obtaining data using a well-controlled experimental protocol in a lab from a relatively small number of participants who viewed television ads, we effectively explained the real-world advertising elasticities of these ads. Although not the main goal of the research, we found that traditional measures explain the most variance in advertising elasticities after controlling for firm differences. This finding gives further support to more than 50 years of advertising research demonstrating that measures such as purchase intent are good predictors of advertising success.

More importantly, we show that the predictions of advertising success can be substantially improved with neurophysiological measures, particularly fMRI, which explained the most incremental variance in advertising elasticities beyond traditional measures. Only one other published study, to our knowledge, has shown such a relationship between neurophysiological measures and market outcomes by using fMRI responses to song clips in the lab to explain subsequent sales of music albums (Berns and Moore 2012). The additional predictive power in our study can be traced back to specific neurophysiological processes (activation in ventral striatum), which tap into a specific construct (desirability). The ventral striatum, through its strong dopaminergic connections, has been shown to play an important role in reward processing. Specifically, it has been associated with the motivation of "wanting" something, rather than just "liking" (Knutson et al. 2007). Therefore, the finding that the ventral striatum explains the most incremental variance over traditional measures in this study is consistent with its role in measuring desirability for the products featured in the ad. Given that advertising firms spend millions of dollars on advertising, our findings have important implications for practice in that they help elucidate which particular methods and exact measures better predict real-world advertising success.

Limitations and Suggestions for Further Research

The pioneering nature of this study opens up possibilities for further research. First, due to the small number of ads, we had to restrict the number of measures in our prediction analysis. Although we were able to use multiple measures across the most common neurophysiological methods, other measures (e.g., pupil dilation) and methods (e.g., facial electromyography, facial coding) were not part of this study. The lower p-values and multiple second-level models could also raise concerns about false positives. However, all second-stage analyses were theoretically motivated and grouped by methodologies to assess additional variation explained beyond traditional measures. We find that fMRI explains significant additional variation (at just over 5%) even when using conservative Bonferroni correction for multiple second-level models. Still, further research should include additional methods, measures, and ads to provide greater degrees of freedom for more comprehensive testing.

Second, we could potentially obtain more precise estimates of ad elasticities by running ads in randomly selected geographic markets to increase the variation in the ad GRP data. One of the challenges in estimating the ad elasticities in this study was that the ads had already been aired and were all part of national campaigns, limiting the variation in GRPs to time-series variation. Directly controlling the variation in ad GRPs would help minimize any biases in estimates of the elasticities resulting from advertising endogeneity. Such endogeneity is not a great concern here because our ad elasticities are consistent with prior literature. Even if there is some bias in our estimates, it should not be correlated with either the traditional or neurophysiological measures, leaving second-stage results unaffected.

Third, we limited all analyses in this study to aggregate data across all 30 seconds of the television ads. However, certain methods, such as biometrics and EEG, may be more effective in identifying interesting variations within portions of the ad (because of their high temporal resolution). These subtle variations may have been washed out when we aggregated our measurement across the entire ad. Therefore, future studies should focus on identifying interesting temporal components within each ad (e.g., branding moments, final seconds) and relate them to advertising success. It is very possible that the biometric and EEG measures may be more effective for these temporal aspects within an ad than measures from fMRI (Ohme et al. 2009).

Finally, relative to traditional methods, neurophysiological methods are typically more expensive and less accessible. Further research could explore the incremental value of each method relative to its cost and accessibility compared with traditional methods.

Conclusion

A wide variety of methods have been developed to assess advertising effectiveness, ranging from traditional self-reported measures to eye tracking and neurophysiological tools. In this study, we provide insight into the relative contribution of each of these methods in the context of television advertising. Specifically, we collected, analyzed, integrated, and compared the role of several methods and measures in predicting real-world advertising success. Our findings clearly demonstrate the potential of neurophysiological measures to complement traditional measures in improving the predictive power of advertising success models. In addition to guiding practitioners toward supplementary measures that could enhance their efforts to predict advertising effectiveness, this study demonstrates the potential of neuroscience applied to marketing research and practice by extending existing measures, helping enrich marketing theories, and improving models of marketing success.

REFERENCES

Ariely, Dan and Gregory S. Berns (2010), "Science and Society Neuromarketing: The Hope and Hype of Neuroimaging in Business," *Nature Reviews Neuroscience*, 11 (4), 284–92.

Barry, Thomas E. and Daniel J. Howard (1990), "A Review and Critique of the Hierarchy of Effects in Advertising," *International Journal of Advertising*, 9 (2), 121–35.

Batra, Rajeev, John G. Myers, and David A. Aaker (1996), *Advertising Management*, 5th ed. Upper Saddle River, NJ: Prentice Hall.

Berns, Gregory S. and Sara E. Moore (2012), "A Neural Predictor of Cultural Popularity," *Journal of Consumer Psychology*, 22 (1), 154–60.

Berridge, Kent C. (2007), "The Debate Over Dopamine's Role in Reward: The Case for Incentive Salience," *Psychopharmacology*, 191 (3), 391–431.

Biel, Alexander L. and Carol A. Bridgwater (1990), "Attributes of Likable Television Commercials," *Journal of Advertising Research*, 30 (3), 38–44.

Brett, Matthew, Jean-Luc Anton, Romain Valabregue, and Jean-Baptiste Poline (2002), "Region of Interest Analysis Using the MarsBar Toolbox for SPM 99," *Neuroimage*, 16 (2), S497.

Brown, Steven P. and Douglas M. Stayman (1992), "Antecedents and Consequences of Attitude Toward the Ad: A Metaanalysis," *Journal of Consumer Research*, 19 (1), 34–51.

Camerer, Colin, George Loewenstein, and Drazen Prelec (2005), "Neuroeconomics: How Neuroscience Can Inform Economics," *Journal of Economic Literature*, 43 (1), 9–64.

Corkin, Suzanne (1984), "Lasting Consequences of Bilateral Medial Temporal Lobectomy: Clinical Course and Experimental Findings in H.M.," *Seminars in Neurology*, 4 (2), 249–59.

Critchley, Hugo D., Pia Rotshtein, Yoko Nagai, John O'Doherty, Christopher J. Mathias, and Raymond J. Dolan (2005), "Activity in the Human Brain Predicting Differential Heart Rate Responses to Emotional Facial Expressions," *Neuroimage*, 24 (3), 751–62.

Davidson, Richard J. (2004), "What Does the Prefrontal Cortex 'Do' in Affect? Perspectives on Frontal EEG Asymmetry Research," *Biological Psychology*, 67 (1/2), 219–33.

———, Paul Ekman, Clifford D. Saron, Joseph A. Senulis, and Wallace V. Friesen (1990), "Approach-Withdrawal and Cerebral Asymmetry: Emotional Expression and Brain Physiology. I," *Journal of Personality and Social Psychology*, 58 (2), 330–41.

———, Klaus R. Scherer, and H. Hill Goldsmith (2009), *Handbook of Affective Sciences*. New York: Oxford University Press.

Delorme, Arnaud and Scott Makeig (2004), "EEGLAB: An Open Source Toolbox for Analysis of Single-Trial EEG Dynamics Including Independent Component Analysis," *Journal of Neuroscience Methods*, 134 (1), 9–21.

Demaree, Heath A., D. Erik Everhart, Eric A. Youngstrom, and David W. Harrison (2005), "Brain Lateralization of Emotional Processing: Historical Roots and a Future Incorporating 'Dominance,'" *Behavioral and Cognitive Neuroscience Reviews*, 4 (1), 3–20.

Dimofte, Claudiu V. (2010), "Implicit Measures of Consumer Cognition: A Review," *Psychology & Marketing*, 27 (10), 921–37.

Dimoka, Angelika (2010), "What Does the Brain Tell Us About Trust and Distrust? Evidence from a Functional Neuroimaging Study," *MIS Quarterly*, 34 (2), 373–96.

——— (2012), "How to Conduct a Functional Magnetic Resonance (fMRI) Study in Social Science Research," *MIS Quarterly*, 36 (3), 811–40.

———, Paul A. Pavlou, and Fred D. Davis (2011), "NeuroIS: The Potential of Cognitive Neuroscience for Information Systems Research," *Information Systems Research*, 22 (4), 687–702.

Du Plessis, Erik (1994), "Recognition Versus Recall," *Journal of Advertising Research*, 34 (3), 75–91.

Foxe, John J. and Adam C. Snyder (2011), "The Role of Alpha-Band Brain Oscillations as a Sensory Suppression Mechanism During Selective Attention," *Frontiers in Psychology*, 2 (accessed April 6, 2015), [available at http://journal.frontiersin.org/article/10.3389/fpsyg.2011.00154/full].

Friston, Karl J., Andrew P. Holmes, Keith J. Worsley, Jean-Baptiste P. Poline, Chris D. Frith, and Richard S.J. Frackowiak (1994), "Statistical Parametric Maps in Functional Imaging: A General Linear Approach," *Human Brain Mapping*, 2 (4), 189–210.

Glascher, Jan and Ralph Adolphs (2003), "Processing of the Arousal of Subliminal and Supraliminal Emotional Stimuli by the Human Amygdala," *Journal of Neuroscience*, 23 (32), 10274–82.

Greenwald, Anthony G., Debbie E. McGhee, and Jordan L.K. Schwartz (1998), "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test," *Journal of Personality and Social Psychology*, 74 (6), 1464–80.

Haley, Russell I. and Allan L. Baldinger (2000), "The ARF Copy Research Validity Project," *Journal of Advertising Research*, 40 (6), 114–35.

Hare, Todd A., Jonathan Malmaud, and Antonio Rangel (2011), "Focusing Attention on the Health Aspects of Foods Changes Value Signals in vmPFC and Improves Dietary Choice," *Journal of Neuroscience*, 31 (30), 11077–87.

Harmon-Jones, Eddie, Philip A. Gable, and Carly K. Peterson (2010), "The Role of Asymmetric Frontal Cortical Activity in Emotion-Related Phenomena: A Review and Update," *Biological Psychology*, 84 (3), 451–62.

———, Lacey Lueck, Meghan Fearn, and Cindy Harmon-Jones (2006), "The Effect of Personal Relevance and Approach-Related Action Expectation on Relative Left Frontal Cortical Activity," *Psychological Science*, 17 (5), 434–40.

Hess, Eckhard H. and James M. Polt (1960), "Pupil Size as Related to Interest Value of Visual Stimuli," *Science*, 132 (3423), 349–50.

Horstmann, Nina, Andrea Ahlgrimm, and Andreas Glockner (2009), "How Distinct Are Intuition and Deliberation? An Eye-Tracking Analysis of Instruction-Induced Decision Models," *Judgment and Decision Making*, 4 (5), 335–54.

Huettel, Scott A., Allen W. Song, and Gregory McCarthy (2008), *Functional Magnetic Resonance Imaging*, 2nd ed. Sunderland, MA: Sinauer Associates.

Jensen, Ole and Ali Mazaheri (2010), "Shaping Functional Architecture by Oscillatory Alpha Activity: Gating by Inhibition," *Frontiers in Human Neuroscience*, 4 (accessed April 6, 2015), [available at http://journal.frontiersin.org/article/10.3389/fnhum.2010.00186/full].

Knutson, Brian, Mauricio R. Delgado, and Paul E.M. Phillips (2010), "Representation of Subjective Value in the Striatum," in *Neuroeconomics: Decision Making and the Brain*, Paul W. Glimcher, Colin F. Camerer, Ernst Fehr, and Russell A. Poldrack, eds. San Diego: Academic Press, 389–406.

——, Scott Rick, G. Elliott Wimmer, Drazen Prelec, and George Loewenstein (2007), "Neural Predictors of Purchases," *Neuron*, 53 (1), 147–56.

Lang, Annie, Paul Bolls, Robert F. Potter, and Karlynn Kawahara (1999), "The Effects of Production Pacing and Arousing Content on the Information Processing of Television Messages," *Journal of Broadcasting & Electronic Media*, 43 (4), 451–75.

Leven, Wilfried (1991), *Blickverhalten von Konsumenten: Grundlagen, Messung und Anwendung in der Werbeforschung*. Heidelberg: Physica-Verlag.

Levy, Irfat, Stephanie C. Lazzaro, Robb B. Rutledge, and Paul W. Glimcher (2011), "Choice from Non-Choice: Predicting Consumer Preferences from Blood Oxygenation Level–Dependent Signals Obtained During Passive Viewing," *Journal of Neuroscience*, 31 (1), 118–25.

Lucas, Darrell B. and Steuart H. Britt (1963), *Measuring Advertising Effectiveness*. New York: McGraw-Hill.

Mandler, George (1980), "Recognizing: The Judgment of Previous Occurrence," *Psychological Review*, 87 (3), 252–71.

McClure, Samuel M., Jian Li, Damon Tomlin, Kim S. Cypert, Latane M. Montague, and P. Read Montague (2004), "Neural Correlates of Behavioral Preference for Culturally Familiar Drinks," *Neuron*, 44 (2), 379–87.

Micu, Anca Cristina and Joseph T. Plummer (2010), "Measurable Emotions: How Television Ads Really Work," *Journal of Advertising Research*, 50 (2), 137–53.

Miller, Earl K. and Jonathan D. Cohen (2001), "An Integrative Theory of Prefrontal Cortex Function," *Annual Review of Neuroscience*, 24, 167–202.

Mognon, Andrea, Jorge Jovicich, Lorenzo Bruzzone, and Marco Buiatti (2011), "ADJUST: An Automatic EEG Artifact Detector Based on the Joint Use of Spatial and Temporal Features," *Psychophysiology*, 48 (2), 229–40.

Morwitz, Vicki G. and Gavan J. Fitzsimons (2004), "The Mere-Measurement Effect: Why Does Measuring Intentions Change Actual Behavior?" *Journal of Consumer Psychology*, 14 (1/2), 64–74.

——, Joel H. Steckel, and Alok Gupta (2007), "When Do Purchase Intentions Predict Sales?" *International Journal of Forecasting*, 23 (3), 347–64.

Murphy, Sheila T. and Robert B. Zajonc (1993), "Affect, Cognition, and Awareness: Affective Priming with Optimal and Suboptimal Stimulus Exposures," *Journal of Personality and Social Psychology*, 64 (5), 723–39.

Ogawa, Sonoko, Toni M. Lee, A.R. Kay, and David W. Tank (1990), "Brain Magnetic Resonance Imaging with Contrast Dependent on Blood Oxygenation," *Proceedings of the National Academy of Sciences in the United States of America*, 87 (24), 9868–72.

Ohme, Rafal, Dorota Reykowska, Dawid Wiener, and Anna Choromanska (2009), "Analysis of Neurophysiological Reactions to Advertising Stimuli by Means of EEG and Galvanic Skin Response Measures," *Journal of Neuroscience, Psychology, and Economics*, 2 (1), 21–31.

——, ——, ——, and —— (2010), "Application of Frontal EEG Asymmetry to Advertising Research," *Journal of Economic Psychology*, 31 (5), 785–93.

Pessoa, Luiz and Ralph Adolphs (2010), "Emotion Processing and the Amygdala: From a 'Low Road' to 'Many Roads' of Evaluating Biological Significance," *National Review of Neuroscience*, 11 (11), 773–83.

Phelps, Elizabeth A. (2004), "Human Emotion and Memory: Interactions of the Amygdala and Hippocampal Complex," *Current Opinion in Neurobiology*, 14 (2), 198–202.

Plassmann, Hilke, John O'Doherty, and Antonio Rangel (2007), "Orbitofrontal Cortex Encodes Willingness to Pay in Everyday Economic Transactions," *Journal of Neuroscience*, 27 (37), 9984–88.

——, Thomas Zoëga Ramsøy, and Milica Milosavljevic (2012), "Branding the Brain: A Critical Review and Outlook," *Journal of Consumer Psychology*, 22 (1), 18–36.

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

Potter, Robert F. and Paul David Bolls (2012), *Psychophysiological Measurement and Meaning: Cognitive and Emotional Processing of Media*. New York: Routledge.

Preacher, Kristopher J. and Andrew F. Hayes (2004), "SPSS and SAS Procedures for Estimating Indirect Effects in Simple Mediation Models," *Behavior Research Methods, Instruments, & Computers*, 36 (4), 717–31.

Sabatinelli, Dean, Margaret M. Bradley, Jeffrey R. Fitzsimmons, and Peter J. Lang (2005), "Parallel Amygdala and Inferotemporal Activation Reflect Emotional Intensity and Fear Relevance," *Neuroimage*, 24 (4), 1265–70.

Schlinger, Mary Jane (1979), "A Profile of Responses to Commercials," *Journal of Advertising Research*, 19 (2), 37–46.

Sethuraman, Raj, Gerard J. Tellis, and Richard A. Briesch (2011), "How Well Does Advertising Work? Generalizations from Meta-Analysis of Brand Advertising Elasticities," *Journal of Marketing Research*, 48 (June), 457–71.

Shaw, John C. (2003), *The Brain's Alpha Rhythms and the Mind: A Review of Classical and Modern Studies of the Alpha Rhythm Component of the Electroencephalogram with Commentaries on Associated Neuroscience and Neuropsychology*, 1st ed. Amsterdam: Elsevier.

Singh, Surendra N., Michael L. Rothschild, and Gilbert A. Churchill (1988), "Recognition Versus Recall as Measures of Television Commercial Forgetting," *Journal of Marketing Research*, 25 (February), 72–80.

Smidts, Ale, Ming Hsu, Alan G. Sanfey, Maarten A.S. Boksem, Richard B. Ebstein, Scott A. Huettel, et al. (2014), "Advancing Consumer Neuroscience," *Marketing Letters*, 25 (3), 1–11.

Smit, Edith G., Lex Van Meurs, and Peter C. Neijens (2006), "Effects of Advertising Likeability: A 10-Year Perspective," *Journal of Advertising Research*, 46 (1), 73–83.

Stewart, David W. (1984), "Physiological Measurement of Advertising Effects," *Psychology and Marketing*, 1 (1), 43–48.

Stipp, Horst and Robert P. Woodard (2011), "Uncovering Emotion: Using Neuromarketing to Increase Ad Effectiveness," research report, Advertising Research Foundation, (accessed April 6, 2015), [available at https://thearf-org-aux-assets.s3.amazonaws.com/research/ARF_Using_Neuromarketing_to_Increase_Ad_Effectiveness.pdf].

Strong, Edward K., Jr. (1925), "Theories of Selling," *Journal of Applied Psychology*, 9 (1), 75–86.

Sutton, Steven K. and Richard J. Davidson (1997), "Prefrontal Brain Asymmetry: A Biological Substrate of Behavioral Approach and Inhibition Systems," *Psychological Science*, 8 (3), 204–210.

Tomarken, Andrew J., Richard J. Davidson, and Jeffrey B. Henriques (1990), "Resting Frontal Brain Asymmetry Predicts Affective Responses to Films," *Journal of Personality and Social Psychology*, 59 (4), 791–801.

Venkatraman, Vinod, John A. Clithero, Gavan J. Fitzsimons, and Scott A. Huettel (2012), "New Scanner Data for Brand Marketers: How Neuroscience Can Help Better Understand Differences in Brand Preferences," *Journal of Consumer Psychology*, 22 (1), 143–53.

———, John W. Payne, and Scott A. Huettel (2014), "An Overall Probability of Winning Heuristic for Complex Risky Decisions: Choice and Eye Fixation Evidence," *Organizational Behavior and Human Decision Processes*, 125 (2), 73–87.

Walker, David and Tony M. Dubitsky (1994), "Why Liking Matters," *Journal of Advertising Research*, 34 (3), 9–18.

Wang, Yong J. and Michael S. Minor (2008), "Validity, Reliability, and Applicability of Psychophysiological Techniques in Marketing Research," *Psychology & Marketing*, 25 (2), 197–232.

Wang, Zheng, Annie N. Lang, and Jerome R. Busemeyer (2011), "Motivational Processing and Choice Behavior During Television Viewing: An Integrative Dynamic Approach," *Journal of Communication*, 61 (1), 71–93.

Wedel, Michel and Rik Pieters (2008), "A Review of Eye-Tracking Research in Marketing," in *Review of Marketing Research*, Vol. 4, Naresh K. Malhotra, ed. Armonk, NY: M.E. Sharpe.

Yoon, Carolyn, Richard Gonzalez, Antoine Bechara, Gregory S. Berns, Alain A. Dagher, Laurette Dubé, et al. (2012), "Decision Neuroscience and Consumer Decision Making," *Marketing Letters*, 23 (2), 473–85.

Zola-Morgan, Stuart and Larry R. Squire (1993), "Neuroanatomy of Memory," *Annual Review of Neuroscience*, 16, 547–63.