nielsen

AN UNCOMMON SENSE
OF THE CONSUMER™

BRAIN ACTIVITY PREDICTS SOCIAL TV ENGAGEMENT

NEURO-TWITTER TV LINKAGE STUDY

NIELSEN NEURO,
R&D DEPARTMENT
NIELSEN SOCIAL
2015

BACKGROUND

In today's ever-more-digital environment, social media (Twitter, Facebook, etc.) is an integral part of our lives, shaping how we interact, share ideas, follow news and form opinions, including consumer preferences.

More specifically around television, social activity supports the industry in capturing the social zeitgeist of TV viewership. In fact, in the last several years, networks have started using show-specific hashtags, Twitter promotions, and show-driven Twitter engagement (i.e., actors and creators posting Tweets in real time during shows or promoting behind-the scenes content) to better consolidate and integrate Twitter conversation.

Both TV viewership and Twitter engagement are a reflection of TV programming itself–there are successful shows, which are more likely to engage the viewer, and there are mediocre or disappointing shows that fail to draw in the audiences. What remains unexplored, however, is the mechanism by which events in the show translate into changes in viewership and Twitter engagement. Why do certain TV programs compel viewers to share their thoughts and impressions with others, while other programs do not?

To that end, this study by Nielsen explores how brain activity could be used to predict Twitter engagement during TV programming. Nielsen Social measures conversation on Twitter for every program aired across over 250 U.S. television networks. The granularity and accessibility of Twitter TV data through Nielsen Social enabled this study to focus on the relationship between Twitter TV activity and brain activity. Brain activity may have a relationship to other forms of to other forms of social activity taking place around TV programming, however Twitter TV data was analyzed for the purposes of this report.

On the level of an individual, the relationship between TV program content and Twitter engagement is mediated by the brain. A viewer continuously processes incoming information (i.e., show content), forming instantaneous impressions, attitudes and emotional reactions in response to events of the show. These subjective reactions, in turn, prompt behavioral responses—expressions of emotion, tuning in or dropping out of viewing, and the desire to share with others (either actively by sending Tweets or posting on social media, or passively by reading relevant chatter online). These psychological processes of information evaluation, formation of emotional and cognitive reactions, and resulting action intent and behavioral output are all implemented by the brain. Thus, in order to understand how TV programming influences viewership and Twitter engagement, we have to understand how the brain translates program content into subjective responses and behavioral outputs.

A few existing academic and industry studies suggest that Twitter engagement increases expression of brain signatures of emotion, attention and memory (Neuro-Insight/Twitter study, 2014) and that brain activity can be used to predict TV viewership for a program (Dmochowski et al., 2013), but there have been no studies that directly examine whether brain activity while watching a TV program can predict Twitter engagement or, stated alternatively, whether Twitter volume can be seen as indicative of more fundamental neural processes that reflect increased emotional engagement, attention and memory activation in response to TV programming.

STUDY AIM

The main goal of this research project was to examine neural processes that can explain how subjective experiences and responses to TV programming translate into Twitter engagement (i.e., desire and action to share personal impressions or express opinions about live TV content). To achieve this aim, we selected a number of 1-hr TV shows (8 unique shows, 9 episodes total) ranging in Twitter and viewership popularity. Participants viewed one of these TV episodes, shown in its entirety post-air, while their brain activity was measured. For each episode, minute-by-minute indices of Emotional Engagement, Memory, Attention, and Overall Effectiveness (Nielsen Neuro proprietary metrics measuring engagement) were extracted and correlated with minute-by-minute changes in Twitter volume (Tweets as measured by Nielsen Social) over the course of the episode.

METHOD

TV EPISODES

Selection criteria for TV shows included in this study were meant to ensure comparable show characteristics. Only 1-hr long prime-time serial shows that were airing new episodes at the time of the analysis were used. Eight shows were chosen to provide a reasonable range (low to high) of Twitter engagement and TV viewership volume as measured by Nielsen (see Table 1). Twitter engagement and TV rating estimations were done on the averages for all episodes aired prior to the selection process. Since recruitment age range for neuro data is limited to adults 21-54 years old (see below), shows for which over 50% of viewers or Twitter authors were younger than 21 years old were excluded.

Out of 8 selected shows, 6 aired on Broadcast TV and 2 shows aired on Cable. Three shows were of the reality/competition genre, four shows were dramas, and one show was a documentary serial show. For one of the shows, Program 8, 2 episodes were used, as the original episode selected was aired 2 hours later on the East Coast due to a delay in sports programming shown prior to the episode. By chance, the second episode of Program 8 selected for testing was also delayed (by 51 mins) on the East Coast for a similar reason.

Each episode was recorded in high definition as it aired in real time and was edited *post facto* to remove commercial breaks. Episodes were selected to allow participant data collection within 3-4 days after the episode first aired on TV. All episodes used in this study came from the second half of their respective shows, which made it easier to recruit participants who were familiar with the show and watched it on a regular basis.

SHOW	AVERAGE TWEETS PER MINUTE FOR THIS EPISODE	AVERAGE AUDIENCE PER MINUTE FOR THIS EPISODE
Program 1 Documentary Series	41.4	1,676,054
Program 2 Drama	66.2	910,828
Program 3 Drama	22.0	1,467,356
Program 4 Reality	58.1	1,203,370
Program 5 Drama	118.8	982,500
Program 6 Reality	101.7	2,414,236
Program 7 Drama	30.22	1,402,931
Program 8 ep1 Reality	979.0	2,638,343
Program 8 ep2 Reality	484.7	3,216,952

Table 1. TV audience and Twitter volume averages show a good mix of high volume, medium volume, and low volume.

TWITTER ENGAGEMENT

Twitter engagement was measured using Twitter TV Activity (i.e., Tweets) from Nielsen Social collected during both Eastern and Pacific time zone airings of each episode. Tweets by the network were not included, unless these were user Retweets. The number of Tweets was calculated minute-by-minute for the duration of the show, including commercial breaks. Tweets sent during the Eastern and Pacific time zone airings were aggregated together. Eastern time zone Tweets sent during Pacific time zone airing (and vice versa) were counted as well, even though they were sent before or after the actual episode was broadcast at that time zone. Thus, there is a certain amount of noise inherent in the metric, but it is likely to be comparable across different episodes.

PARTICIPANTS AND PROCEDURES

Participants (21-54 years old) were recruited from the San Francisco Bay Area, Chicago and Atlanta using standard Nielsen Neuro recruitment procedures, which include a representative race and ethnicity sampling consistent with local demographic composition. All participants were screened for neurological or medical disorders or conditions that are known to affect EEG or cognitive functioning. An equal number of males and females were used in each recording cell. An age limit was applied to minimize inter-subject variability in EEG data (still-developing brain of young adults and age-related changes in old adults are known to alter both frequency and amplitude characteristics of EEG signal).

Data for each episode was collected from 36 participants (over 300 participants in total) who indicated that they regularly watched the given show. Participants included people who used Twitter (sent or viewed Tweets) often, sometimes and not at all.

Participants were recruited prior to the episode airing on TV and were asked not to watch this particular episode before the experimental session, which was scheduled 2-4 days after the original air date. During the session, participants were asked whether they did watch the episode, and fewer than 10% indicated that they did. Thus, over 90% of participants were exposed to each episode for the first time. During the session, each participant was seated in a comfortable chair in a soundproof room with comfortable lighting. Episodes were presented without interruption on a large TV screen.

NEURO MEASURES

Continuous EEG was recorded from each subject. Vertical and horizontal eye movements were recorded using external electrodes positioned directly above and to the side of the left eye. Signals were digitized and processed offline. Muscle noise and eye movement artifacts were corrected using in-house Matlab-based component analysis software. Spectral analyses, examining power in several pre-determined frequency bands, were used to derive metrics for Emotional Engagement, Attention, Memory and Overall Effectiveness using Nielsen Neuro proprietary algorithms. For each metric, spectral data was normalized within each subject, averaged for the group on the minute-by-minute basis, and fit to a normalized 10-point scale. Neuro data was obtained for TV programming only (not for commercial breaks).

DATA PREPROCESSING AND ANALYSIS

EPISODE CHARACTERIZATION

The Tweets about each episode were subjected to word cloud analysis to assess the nature of Twitter engagement for that episode. Show-related information was also gathered for each episode. The results confirmed that shows with high Twitter engagement generated many diverse comments relevant to the content of the episode.

In contrast, Twitter content for shows with low Twitter engagement was dominated by few users reposting promotional Tweets, most of which were not related to the content of the episode.

Only 2 episodes showed low Twitter engagement–Program 3 and Program 7, with other episodes generating Twitter content closely linked to events of the show.

Episode analysis revealed that the Program 2 episode tested was the last episode before the season finale and featured several big reveals and death of one of the key characters. The Program 5 episode was the season finale. The Program 6 episode featured two competitions and elimination. The two Program 8 episodes aired with a delay on the East Coast, and prior to the second episode, one of the competitors ran a promotional campaign, resulting in more than 20,000 Retweets.

TWITTER DATA PREPROCESSING

Given the range in Twitter volume across different shows (Figure 1), raw minute-by-minute Twitter values do not provide sufficient separation among the shows, especially for episodes where Twitter volume was less than 500 Tweets per minute. To normalize the distribution of Twitter engagement values, we applied logarithmic and square-root transforms, with square root providing the best normalization.

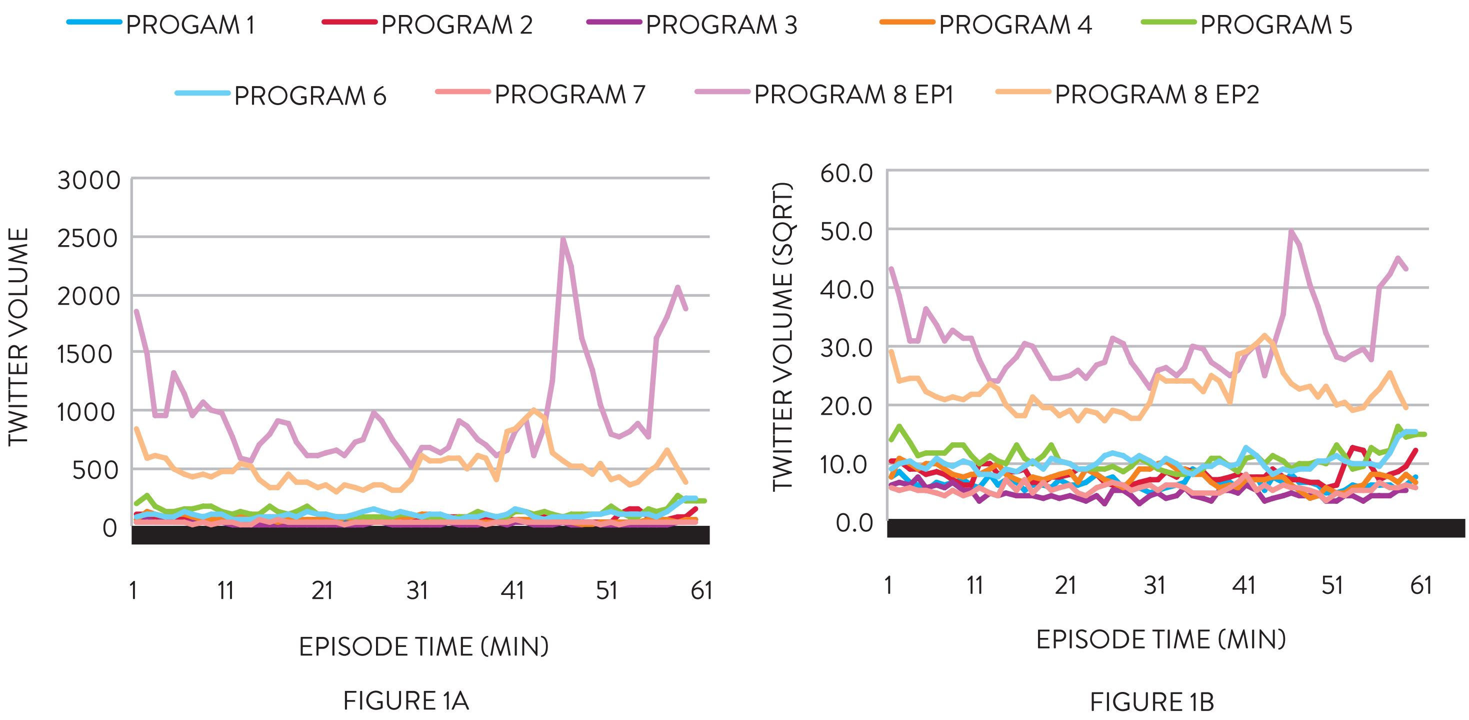

Figure 1. Twitter volume minute-by-minute data for each episode (left panel – raw volume, right panel – square root normalized volume data)

TIME LAG ADJUSTMENT

We hypothesized that there would be a lag between the neuro and Twitter data streams. In addition to the time it takes for cortical areas to process incoming information (usually around 100 ms) and, thus, for EEG signal to reflect brain activity associated with episode content, there is also a lag in time before changes in brain activity translate into action (i.e., time that it takes to send a Tweet about the show). We assessed 3 possible lags between neuro and Twitter data—0 minutes, 1 minute, and 2 minutes. Since different programs could have different patterns of engagement (depending on the structure and content of each episode), we customized the lag for each program, selecting the lag that provided the best fit between minute-by-minute neuro and Twitter data.

SEGMENT ANALYSIS

To assess the relationship between brain responses to TV episodes and Twitter engagement across all programs, we collapsed minute-by-minute data within each segment (a continuous programming epoch between commercial breaks). This level of analysis allowed us to smooth over random minute-by-minute fluctuations in data, providing a more robust estimate of both neural and Twitter activity. Time-adjusted Twitter and neuro data were normalized on the minute-by-minute basis across all programs (z-score transformation) before they were averaged within each segment, resulting in 49 total segments. Since Twitter and neuro data fluctuate at different scales, we hypothesized that the rate of change (i.e., minute-by-minute difference in neuro and Twitter values respectively) would be a more appropriate metric than the raw data. Analyses confirmed this by showing stronger correlations for the rate of change than for raw values.

COMMERCIAL BREAK ALLOCATION

Further analysis of Twitter engagement patterns revealed that for most shows increases in Twitter volume occur during commercial breaks (i.e., people tend to send Tweets about what happened during a previous episode segment while watching commercials immediately following that segment; Figure 2).

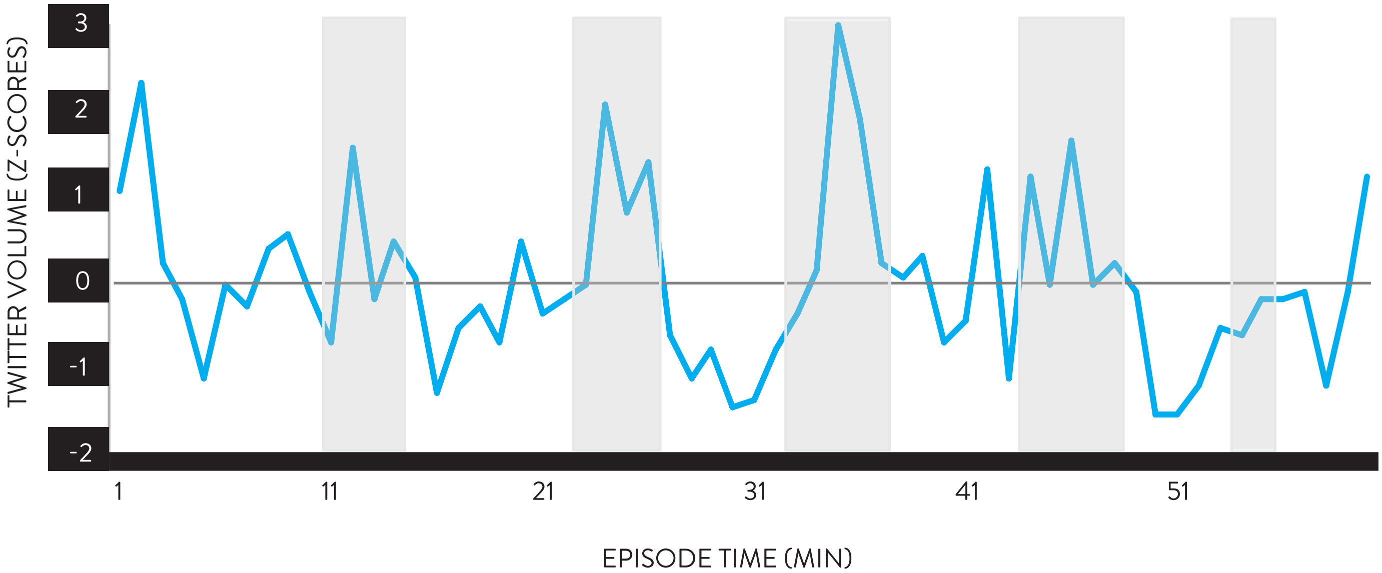

FIGURE 2

Figure 2. Twitter volume (normalized) for a sample episode. Gray shaded areas indicate commercial break times.

This finding indicates that Twitter volume during commercial breaks represents the remnant impact from the preceding program segment. To account for this temporal shift, it is necessary to reallocate Tweets from ad breaks into the program segments. After analysis of optimal ad break reallocation, we found that reallocating the first two minutes of a commercial break to the Twitter values for the previous segment yielded the best correlations.

MANAGING EXTREMES AT THE BEGINNING AND END OF EPISODES

We observed that Twitter data considerably fluctuated at the beginning and at the end of tested episodes, often resulting in extreme/outlier values. Indeed, Twitter engagement during the first few minutes of the episode is likely to reflect overall excitement and Twitter chatter about the upcoming episode (i.e., not content-specific). Similarly, at the end of the show, Twitter volume reflected people reacting to the whole show rather than the previous segment alone, which also results in extreme values. Given that these extreme values skewed overall segment data for Twitter volume, we chose to omit these values from analyses, optimizing omission periods (1-3 minutes) for each program.

RESULTS

The best prediction score between Twitter and neuro data was seen when the three primary neurometrics–Emotional Engagement, Memory and Attention–were combined as a sum (omitting any of the metrics considerably reduced the correlation). Using optimized Twitter-to-neuro lag and optimized omission period at the beginning and the end of each program, we observed a strong positive correlation ($r=0.795$, $r2=0.631$) between Twitter volume and EEG measures of Emotion, Memory, and Attention (Figure 3). When the lag alone was optimized (and not the omission period), the r value was 0.733.

TWITTER TV ACTIVITY VS. NEUROLOGICAL ENGAGEMENT (MINUTE-BY-MINUTE RATE OF CHANGE)

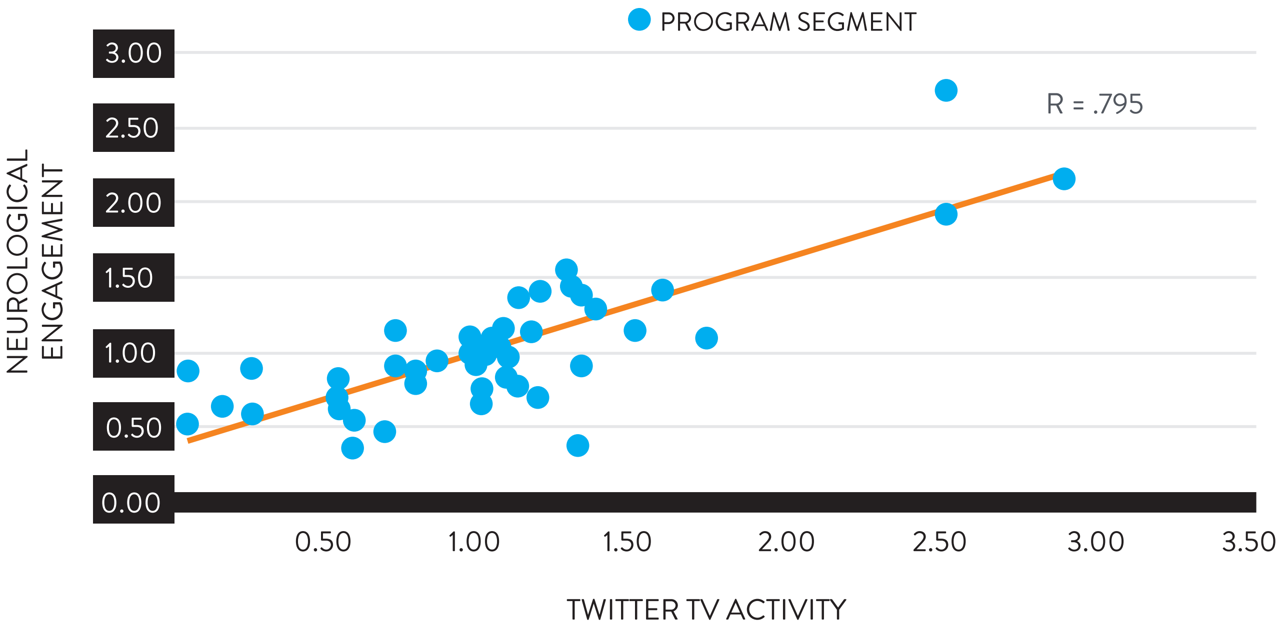

Figure 3. Correlation between normalized Twitter volume scores and combined neurometrics (Emotion+Memory+Attention).

Source Nielsen: 7/31/14-9/14/14. Segment-level analysis of eight prime time broadcast and cable TV programs (9 episodes) ranging in levels of Twitter activity and TV ratings. Neurological Engagement was measured as the sum of indices of Emotion, Memory, and Attention. Twitter TV activity was measured as relevant U.S. Tweets. Normalizing transformations were used to smooth the Twitter TV and neuro data for modeling purposes.

These results provide a mechanism by which TV content is translated into program-related social engagement (i.e., Twitter volume). Specifically, segments that elicited strong brain responses were also associated with high Twitter engagement. It is particularly important to note that the relationship between Twitter and brain activity was strongest when EEG metrics for Attention, Emotion and Memory were used in combination, suggesting that the overall *salience* of content is the driving force in Twitter engagement (i.e., people are more likely to be compelled to send Tweets about their experiences and impressions when the program content recruits multiple psychological processes, effecting change in both subjective experience and brain activity across all levels). Thus, an increase in Twitter volume can be seen as directly indicative of greater cognitive and neural engagement of viewers with the program.

There are three major implications from this study:

- With brain activity predicting social response, networks and content producers can use neurological testing separate of or as a complement to existing testing practices to optimize programming.
- TV networks can view Twitter TV activity around a program's live airing as a bellwether for understanding how engaged TV audiences are with programming (overall, and minute-by-minute as programs unfold).
- Agencies and advertisers can look to Twitter TV metrics as a part of the media planning and buying process to identify shows with engaged audiences and, by extension, opportunities to potentially increase ad memorability and sales outcomes.

FUTURE DIRECTIONS

Now that we have established that brain activity reflecting changes in Attention, Emotional Engagement and Memory can predict Twitter engagement across multiple TV programs, it is a natural next step to examine whether neural activity can also predict changes in TV viewership. Our hypothesis is that neural data can be used to identify segments of TV programming that do not resonate well with viewers, exhibiting low Emotion, Memory or Attention scores, and at those moments, viewers will be more likely to drop out. On the other hand, salient segments can prompt an increase in viewership by increasing the probability that a casual viewer switching channels will stay and continue to watch the program or a viewer seeing other real-time episode-related activity (online or otherwise) will tune in. Similarly, another step in this line of research is to examine whether ad performance changes depending on whether it is preceded by a segment with high neural and Twitter engagement or by a segment that did not resonate with viewers.

In combination, these results can have important implications for TV programming development. If neurometrics can help identify "bad" and "good" segments of the program that will translate into high or low Twitter engagement and viewership increases or drop outs, this can be an invaluable tool for early program testing, providing key information to editors, developers and producers of TV content. Neuro testing can be done early in the deployment process, separate of or in addition to existing testing practices, to identify shows that are likely to be successful with viewers or to suggest areas for improvement at the level of specific episodes. For existing programs, Twitter volume can be used to indicate how engaging the programs are and what segments resonated most strongly with viewers.

ABOUT NIELSEN

Nielsen N.V. (NYSE: NLSN) is a global performance management company that provides a comprehensive understanding of what consumers Watch and Buy. Nielsen's Watch segment provides media and advertising clients with Total Audience measurement services across all devices where content — video, audio and text — is consumed. The Buy segment offers consumer packaged goods manufacturers and retailers the industry's only global view of retail performance measurement. By integrating information from its Watch and Buy segments and other data sources, Nielsen provides its clients with both world-class measurement as well as analytics that help improve performance. Nielsen, an S&P 500 company, has operations in over 100 countries that cover more than 90 percent of the world's population.

For more information, visit www.nielsen.com.

nielsen AN UNCOMMON SENSE OF THE CONSUMER™

nielsen

AN UNCOMMON SENSE
OF THE CONSUMER™