Zombies, Brains, and Tweets: The Neural and Emotional Correlates of Social Media

Brian Abelson

Harmony Institute

January 11, 2013

1 Introduction

What constitutes audience engagement? What elements of a show produce the most social activity? Much attention has been paid to Twitter's ability to capture real-time audience reactions and sentiment towards a television show, but little is known about what subjects drive an individual to tweet (or conversely, not to tweet). Though common sense may suggest that especially provocative, humorous, or emotional moments generate the most activity on social media, are these moments also the most neurologically stimulating? By combining social media data, content analysis, and brain scans, this study attempts to answer these questions using AMC's hit show, *The Walking Dead*, as a case study.

2 The Rise of Social Television

Over the past five years, social media has become a powerful platform for capturing audience engagement with entertainment. Televised spectacles like award shows, the Olympics, or the U.S. presidential debates generate hundreds to thousands of mentions on social media per second, providing marketers, advertisers, and content-creators with rich streams of emotional reactions in real-time (Whitaker, 2012; Downing, 2012; Jannarone, 2012). Subsequently, numerous academics have used these data streams to successfully predict box office revenue for films and ratings of television shows (e.g. Subramanyam, 2011; O'Brien, 2010) . Twitter now provides television networks with a set of best practices for "live-tweeting," encouraging producers to develop unique hashtags and have their stars tweet during episode airings (Twitter, 2012). The company has also recently teamed up with Nielsen to create a new metrics for engagement with television (Greenmeier, 2013). The social revolution of television is such that a show's content is often altered in response to public commentary (Jannarone, 2012).

Yet despite this rapid transformation of how people engage and interact with televised content, it is still uncertain whether social media is an accurate indicator of audience engagement. As one network executive recently exclaimed, "I would love to do an episode that was so amazing you got fewer Tweets" (quoted in Jannarone, 2012). To put this assertion in other words, stimulating content may not necessarily encourage sharing. Given the power with which Twitter can seemingly predict the critical and monetary success of entertainment, why do many still contend that social media poorly captures audience engagement?

3 What Do People Share?

One potential reason for this disconnect may be the difference between stimulating and shareable content. Berger and Milkman (2009) investigate this difference using observational and experimental methods. Working with a corpus of 7500 stories on the *New York Times* website over one year, the authors found that articles which evoked "activating" emotions in readers – like surprise and shock – showed up more often on the *Times'* "most emailed" list, even when controlling for a variety of other factors. Conversely, articles that elicited "deactivating" emotions – like sadness – were far less likely to be shared. These findings were then replicated in a variety of experimental settings in which the content of the articles was manipulated to isolate the effects of these emotional characteristics.

While Berger and Milkman focus their study on news media, their findings have clear implications for entertainment. If people are more inclined to share content that elicits certain emotions, then social media responses to movies and television should similarly preference particular narrative moments. However, we believe that findings from research on news media are not so easily applied to televised entertainment. In particular, emotional reactions to nonfictional accounts of events are not necessarily comparable to those of fictional content. Building upon recent popular science literature and research on game design, we replace Berger and Milkman's concept of "activation" with "immersion" – a condition in which a viewer is deeply and personally invested in a narrative, evidenced by an intense, emotional, and even humorous response to a piece of media or content. Best explained by clichés like, "getting lost in a film" or "seeing one's self in a character," we argue that immersion is a better metric of engagement than mere sentiment or activation, which too often ignore the context in which these emotions are expressed.

4 Using Neuroscience to Measure Engagement

To study the relationship between stimulating and sharable entertainment, one must first establish a baseline measurement of audience engagement. Here, recent advances in neuroscientific research on natural vision present a direction forward. Historically, this field has used highly constrained stimuli as it was assumed that complex inputs like film or television would produce highly inconsistent results across subjects. However, Hasson et. al (2004) inverted this approach, choosing instead to allow five subjects to freely view a 30-minute film while being monitored with functional magnetic resonance imaging (fMRI). The researchers then calculated inter-subject correlations throughout the film to isolate "preferred stimuli embedded in the complex stimulation sequence" (1634). What they found is that subjects' brain activity was surprisingly similar throughout the film. As the researchers put it, "despite the completely free viewing of dynamical, complex scenes, individual brains 'tick together' ... when exposed to the same visual environment" (1634). Taking this result a step further, the researchers also found that the scenes that produced the highest levels of synchronicity seemed to consist of content that was "emotionally activating" (1638).

Yet while such results suggest a promising step forward, the use of fMRI to measure audience engagement is significantly constrained by the unnatural setting in which it is administered, its prohibitive costs, and its low temporal resolution. To address these shortcomings, Dmochowski et. al (2012) employ a similar 'reverse-correlation' method using entroencephalography (EEG), allowing for the computation of correlations down to a single second. Similar to Hasson et. al (2004), Dmochowski et. al (2012) find that spikes in correlated brain activity "occur in remarkable correspondence with arousing moments of the film " (1). These correlations decreased upon subsequent viewings of the film and effectively disappeared when the stimulus' temporal sequence was randomly shuffled. The researchers argue that these findings suggest that moments of increased inter-subject correlation represent instances of "emotionally-laden attention" – a phenomenon that might be broadly defined as "engagement."

5 Study Design

5.1 Hypotheses

Drawing upon these related insights from social psychology and neuroscience, we construct a novel study design to compare patterns of social media sharing and neural engagement over the course of a television show. We chose the 90-minute series premiere of *The Walking Dead* – a show on AMC about the zombie apocalypse - as our test case. We selected this program as it generates a high volume of mentions on social media while also comprising varied emotional content. Following the research outlined above, we hypothesize the following:

H1: Social media activity is positively correlated with instersubject neural sychronicity.

While a gap may exist between stimulating and shareable content (as Berger and Milkman (2009) suggest), their study also found that overall interest in a story was an extremely strong indicator of sharing. Following this insight, if neural synchronicity is correlated with engagement, we should expect that moments of the show that generate increased intersubject correlation will also generate increased social media activity.

H2: Content which evokes reactions associated with positive immersion are more strongly correlated with social media activity than positive sentiment.

As we argue above, using sentiment as a metric for engagement is problematic in that it ignores the context in which these emotions are expressed. In turn, we should expect that our more nuanced scale of immersion – which takes into account an audience's level of personal invovlement with a narrative – will serve as a better indicator of engagement than sentiment. In the next section, we provide an overview of our three-pronged methodology combining content analysis, social media analysis, and EEG scans.

5.2 Methods: Towards a Bridge Between Neural and Social Networks

5.2.1 Content Analysis

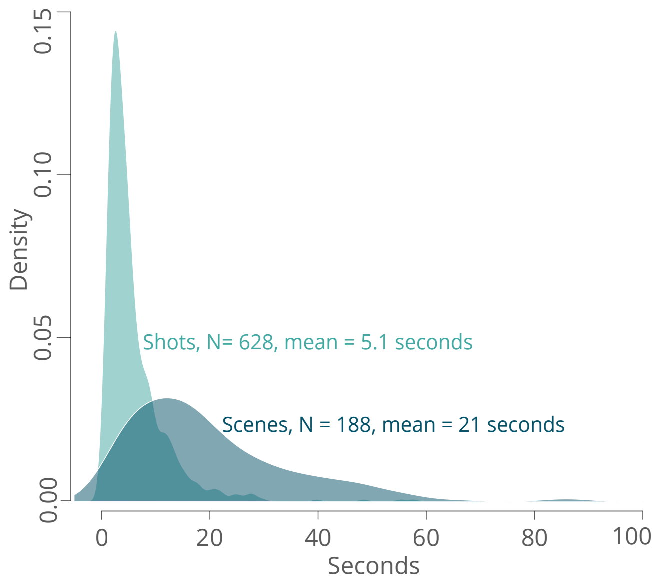
We began by obtaining a copy of "Days Gone By" from iTunes (Darabont and Kirkman, 2010). Using FinalCut Pro, we systematically hand-coded the timestamp of the beginning and end of each of the show's 628 shots. For each of these shots, we also recorded which characters were featured on screen, whether violent acts occurred, and certain cinematographic elements like shot techniques (e.g. wide-angle or close-up) and camera movement. This objective classification of on-screen elements provided us with a highly detailed overview of the show, allowing us to report statistics such as:

- 62% of the shots included the main character, Rick.
- 15% of the shots included zombies.
- 7% of the show featured acts of violence.
- The show depicted 19 gun shots to the head.

Though shots provide an objective standard with which to categorize various moments of the show, when faced with the task of linking the show's content with social media responses, a more capacious narrative unit was needed. To address this issue, we drew upon the idea of a "scene," or a series or related shots that depict an event that advances the story arc. We thus define scenes as an aggregate of shots that constitute a distinct unit of the show's narrative. Through a collaborative back-and-forth process, we divided the show into 188 scenes.

Figure 1 shows the distribution of shot and scene lengths. The majority of shots are less than ten seconds with a mean of 5.1 seconds while scenes follow a slightly more normal distribution centered on a mean of 21 seconds.

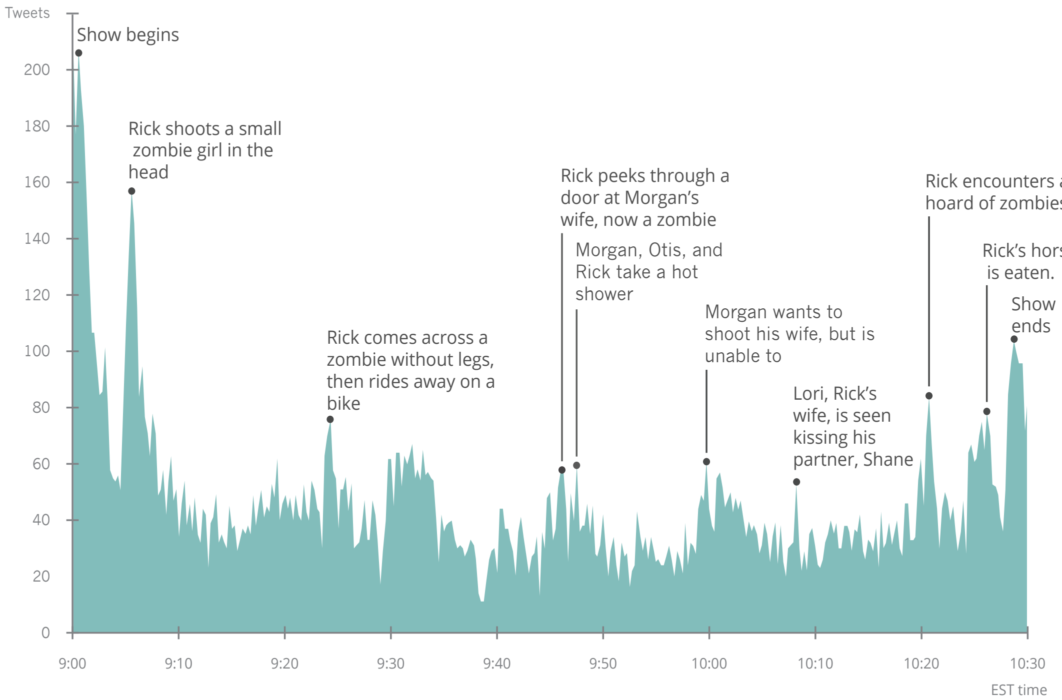
Figure 1: Shot and Scene Lengths



5.2.2 Social Media Analysis

Armed with a detailed dataset of the show's content, we then worked with Crimson Hexagon (a social media analytics platform) and Twitter to obtain each of the 19,000 relevant tweets sent out during the hour-and-a-half series premiere on October 31, 2010. Figure 2 displays these tweets over the course of the show with labels added for moments of the narrative that seemed to elicit spikes in activity. However, not all scenes

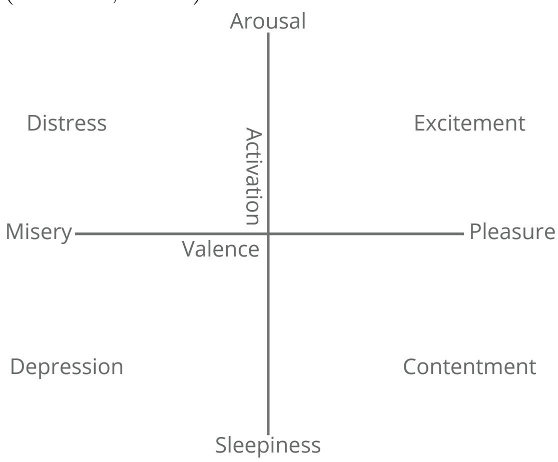
Figure 2: Tweets during the initial airing of *Days Gone By*



could be reliably mapped to such spikes. Furthermore, many users referenced no content at all, simply declaring that they were watching the show or that the show was on. To faithfully link each tweet back to the content it mentioned, we scoured the tweets for those that directly referenced scenes in the show. For instance, we got rid of all tweets that said "Now watching #thewalkingdead" or "#thewalkingdead is great!" This left us with 1947 relevant tweets. We then referenced each of these tweets to any and all scenes mentioned in the message. For instance, general responses like "the opening sequence was awesome! #thewalkingdead" were associated with all 14 scenes in the opening sequence. Highly specific responses like "Boom! Headshot #thewalkingdead" were linked to a single scene. When there was any confusion as to what particular scene the tweet was referencing, we looked at the time the message was sent out and cross-referenced it with what was happening at that time in the show. Finally, to ensure that scenes that elicited vague responses did not receive more weight than those that garnered specific responses, we weighted each tweet by the total number of scenes mentioned in the tweet. As a result, general tweets contributed less to each of the scenes they referenced than those which clearly referred to a single scene.

A second dimension of social media analysis involved the identification of emotions embedded within each tweet. We initially experimented with various methods for automatically classifying the sentiment and/or emotion of text. Unfortunately, these tools were not designed for documents of 140 characters or less and were thus not reliable at the level of an individual tweet. To address this challenge, we developed our own taxonomy for categorizing the emotions of responses to the show on Twitter. We grounded this schema in psychological research which simplifies the spectrum of emotions down to the interplay of two dimensions: valence and activation (Russell, 1980).

Figure 3: A Circumplex Model of Affect (Russell, 1980)

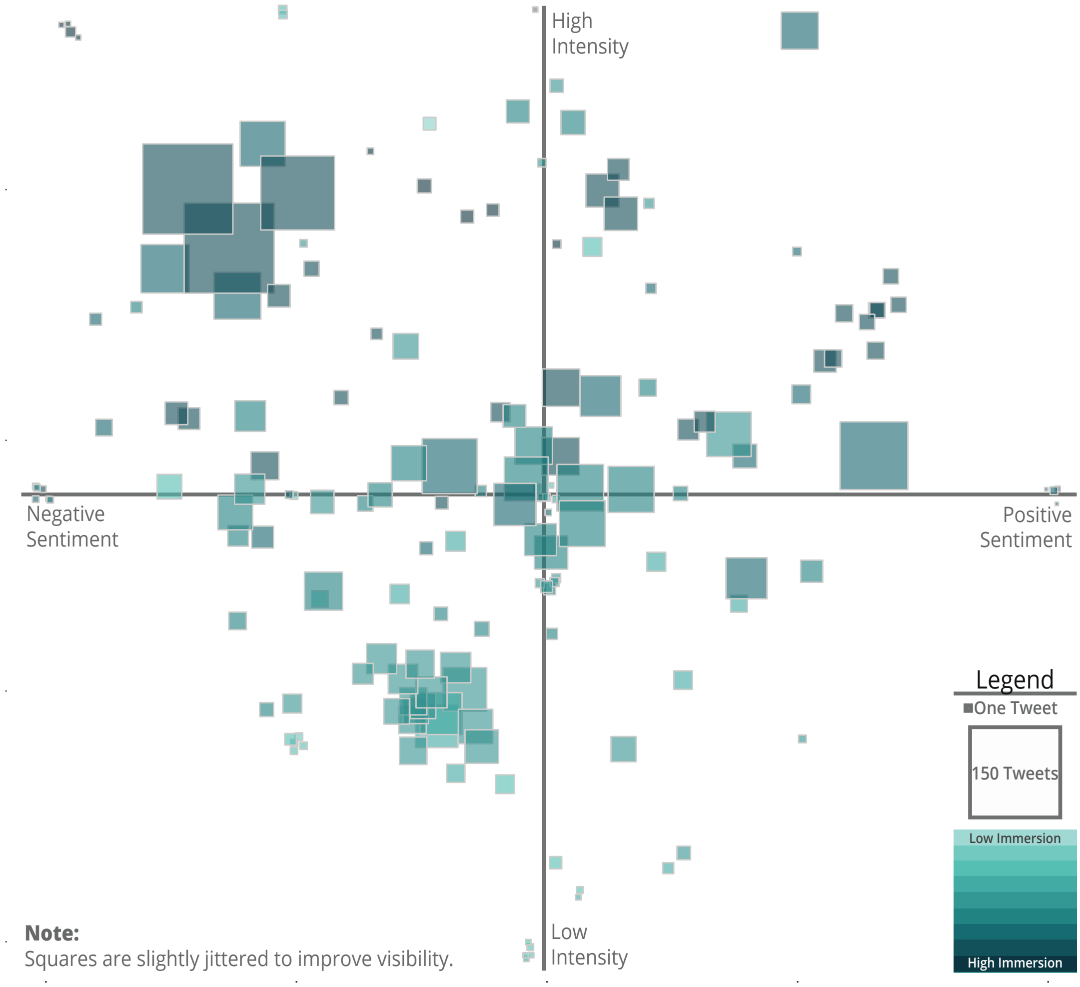


In Figure 3 (a reproduction of a graphy from Russell (1980)), "valence" represents a scale ranging from displeasure to pleasure, or negative to positive sentiment, while "activation" represents a continuum from sleep to arousal, or no activity to intense activity. While this bivariate model imperfectly captures the complexity of human emotion, it does provide a set of relatively objective standards that guided our hand labeling of tweets.

Using this typology as our guide, we coded each tweet on a three-point scale for "sentiment" (-1 = negative, 0 = neutral, 1 = positive) and "intensity" (0 = no intensity, 1 = some intensity, 2 = high intensity). In addition to these two emotional components, we were also interested in a number of other factors: Did the message comment on the actor's abilities or the show's production values? Was the viewer personally affected by the content? Was the tweet intended as humorous? Therefore, we also included binary variables for "personal" (0 = no, 1 = yes), "show" (0=no, 1=yes), and "humor" (0 = no, 1 = yes) in order to approximate a viewer's level of immersion with the subject material referenced. To estimate the level of immersion expressed in each message, we developed an algorithm that weighted all possible combinations of "sentiment", "show", and "personal", giving positive values to emotional comments expressing personal investment in the narrative and negative values to matter-of-fact commentary on the show. These weights are then amplified if the tweet also scored high for intensity or was humorous.

In Figure 4, we visualize the interplay of emotion, immersion, and social media activity at the level of the scene, replicating the layout of Figure 3, with sentiment on the x-axis and intensity on the y-axis. Each square represents a scene in the show. The squares are sized by the number of tweets about the scene and colored by immersion. In this chart we see a cluster of scenes that evoked a high number of intense, negative reactions and also scored high for immersion. This provides preliminary evidence to support our assertion that negative sentiment does not mean a lack of engagement. There also seems to be a clear link between high social media activity (larger squares) and immersion (darker squares). We will expand upon these findings in Section 7.2.

Figure 4: Total Tweets and Sentiment, Intensity, and Immersion per Scene



5.2.3 EEG Scans

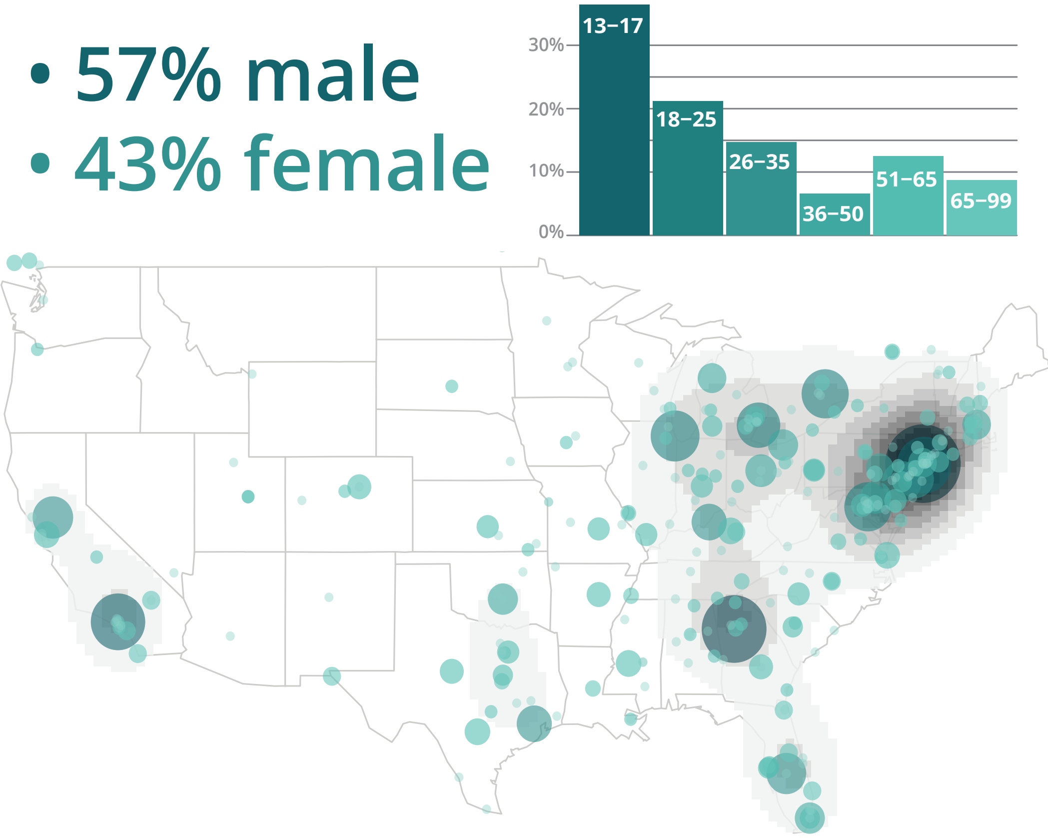
The final component of our methodology entailed the use of EEG data as a proxy for audience engagement. Working with Jacek Dmochowski and Lucas Parra of City University of New York, we rigorously matched the demographics of the individuals in our social media sample to those of our participants. Focusing on the 1,062 unique individuals who sent out 1,947 tweets associated with content in the show, we mined the profiles of these users for their gender, age, and location.

We began by extracting users' first names and matching them with data on baby names listed on the the official website of U.S. Social Security Administration. While this process only matched about 60% of the individuals in our dataset (many people choose to not include their full names in their profiles), an analysis of a random sample of 100 of these individuals suggested that this process was 90% accurate. Next, we exploited the "location" field in profiles to estimate the geographic distribution of the individuals in our dataset. Through a combination of algorithmic matching, hand-cleaning, and geocoding, we were able to successfully match 75% of individual's locations. Finally, to estimate the age of viewers in our dataset, we collected the entirety of each individual's twitter history and fed it into an algorithm on `http://www.uclassify.com/` which uses text features to predict someone's age range.

Our sample comprised mostly young people – the majority of which were male – residing in urban areas on the East Coast (see Figure 5). This was unsurprising given the show's target audience, previous research on the general characteristics of social media users, and the fact that the show aired first on the East Coast due to time zone differences.

Using these demographics as our guide, we attempted to recruit subjects who matched this profile as much as possible. A serious limitation, of course, was geography, and we were unable to recruit participants outside of New York City. Given that our assessment suggested a skew towards urban areas, however, we felt this was an acceptable compromise and would not overly bias our results. Prospective subjects filled out a survey documenting their social media usage, entertainment preferences, and familiarity with the show. Any respondents that did not use social media or had already seen an episode of *The Walking Dead* were

Figure 5: Estimated Viewer Demographics on Twitter for *Days Gone By*



excluded from the study.

20 participants were successfully recruited for data collection. Each was shown the 90-minute episode while wearing a 64-electrode EEG cap. We also recorded the electrooculogram (EOG), or "the difference in electrical charge between the front and back of the eye that is correlated with eyeball movement," so as to remove these components in the post-processing phase. Finally, following standard practice, the signals were filtered at 60 Hz to remove "power noise."

The resulting neural data was operationalized through an equation that maximizes the correlations between all 480 (20 * 19) possible subject pairs. This innovative decomposition method pioneered by Dmockhoski et. al (2011) combines the readings from each of the 64 electrodes to obtain common linear components of neural activity across all subjects. This process reduces the noise inherent to EEG and enables analysis of precise moments of the stimulus that were most engaging. For a more thorough overview of the process of collecting and operationalizing the neural data, please refer to the authors' original paper overviewing their methodology.

6 Model Selection

We address our two hypotheses through regression analysis (Tables 2 – 5). In these models our outcome variable is the number of tweets per scene, with each tweet weighted by the total number of scenes mentioned in the message. A control is added for the duration of each scene to account for the effect of our subjective classification of these narrative moments (see Section 7.1 for more). EEG data is expressed through the three highest correlated components of neural activity throughout the show. The emotions embedded within each tweet are included as outlined above in Section 6.2 and operationalized as per-scene-averages. Finally, controls are added for the presence of shots that feature zombies or violence within a scene. These are included to ensure that any detected effect is not simply the product of content unique to a show about the zombie apocalypse. An overview of these variables is presented in Table 1.

From these summary statistics, it is apparent that the distribution of the outcome variable (in **bold** in Table 1) is highly skewed to the right,

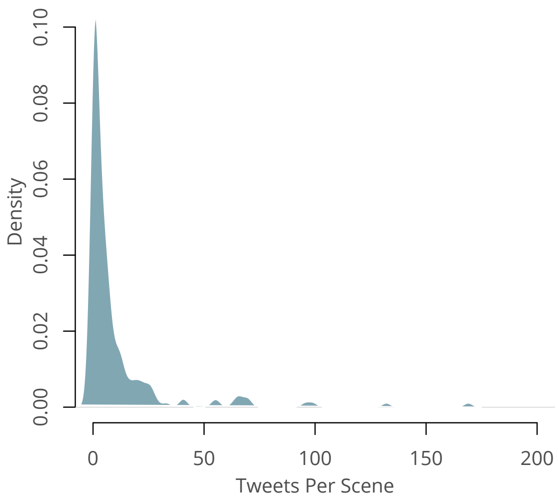
Table 1: Summary Statistics

Variable	N	Min.	Q1	Med.	Mean	Q3	Max.	Std. Dev
tweets per scene	188	0.00	1.80	8.00	23.40	25.20	202.00	36.20
— weighted	**188**	**0.00**	**0.67**	**2.90**	**10.30**	**8.90**	**169.00**	**21.90**
scene duration	188	1.40	9.60	17.20	20.90	28.10	90.80	15.90
intensity*	1947	0.00	1.00	1.00	1.20	2.00	2.00	0.67
— *per scene*	188	0.00	0.55	1.00	0.89	1.30	2.00	0.56
sentiment*	1947	-1.00	-1.00	0.00	-0.21	1.00	1.00	0.83
— *per scene*	188	-1.00	-0.38	0.00	-0.11	0.00	1.00	0.43
show*	1947	0.00	0.00	0.00	0.20	0.00	1.00	0.40
— *per scene*	188	0.00	0.00	0.15	0.25	0.43	1.00	0.29
personal*	1947	0.00	0.00	0.00	0.15	0.00	1.00	0.36
— *per scene*	188	0.00	0.00	0.06	0.09	0.15	0.53	0.12
humor*	1947	0.00	0.00	0.00	0.26	1.00	1.00	0.44
— *per scene*	188	0.00	0.00	0.12	0.18	0.31	1.00	0.21
immersion*	1947	-2.60	-0.73	0.38	0.00	0.76	1.90	1.00
— *per scene*	188	-2.30	-0.52	-0.06	-0.19	0.10	1.20	0.54
zombie(s)	628	0.00	0.00	0.00	0.18	0.00	1.00	0.39
violence	628	0.00	0.00	0.00	0.08	0.00	1.00	0.27
first component	7862	-0.07	-0.01	0.01	0.02	0.05	0.45	0.05
second component	7862	-0.07	-0.02	0.00	0.01	0.03	0.31	0.04
third component	7862	-0.07	-0.02	0.00	0.01	0.03	0.29	0.03

* *included for comparison; regressions use per-scene averages as inputs.*

with a mean of 10.3, a median of 2.9, and a standard deviation of 21.9. This high degree of dispersion (further evidenced by Figure 6) suggests that Negative Binomial Regression is well-suited to the data. All independent variables were scaled from zero to one to allow comparison of effect sizes. Finally, models were fit using the command "bayesglm" in the "arm" pack-

Figure 6: Density of Weighted Tweets



age for R which estimates the parameters of a negative binomial regression. While models were fit in both bayesian and frequentist frameworks, the results were comparable, so we present frequentist parameters to aid interpretability.

6.1 Notes on Nesting in the Data

Given the levels of nesting in the data (188 observations of social media activity per scene, 628 observations of content attributes per shot, 7862 observations (or half-second windows) of EEG data for the show), we might conclude that a multi-level or mixed-effects model is preferred. However, the fact that tweets are operationalized at the level of the scene – the group with the lowest resolution – prevents us from using such methods since the response variable is already calculated on a per-scene rate. This is a serious limitiation to our study as it is not reasonable to assume that windows of neural activity within a given scene are wholly independent.

As a sanity check, fixed effects models were fit with brain activity and tweets nested within 17 sequences – or distinct spatio-temporal moments throughout the show. While these results were similar to the non-nested models presented below, we feel that sequences obfuscate the temporal richness of the neural and social media data, which can vary drastically from observation to obvservation.

Here, we present models with the social media and content covariates replicated across each window of EEG activity within a given scene. We attempt to control for any bias introduced in this process by inserting the duration (in seconds) of each scene. In theory, this added variable should offset the influence of the number of EEG observations per scene on the percieved relationship between brain and social media activity within that scene. Moving forward we hope to improve this methodology so as to more faithfully model the individual and group-level effects of each data source.

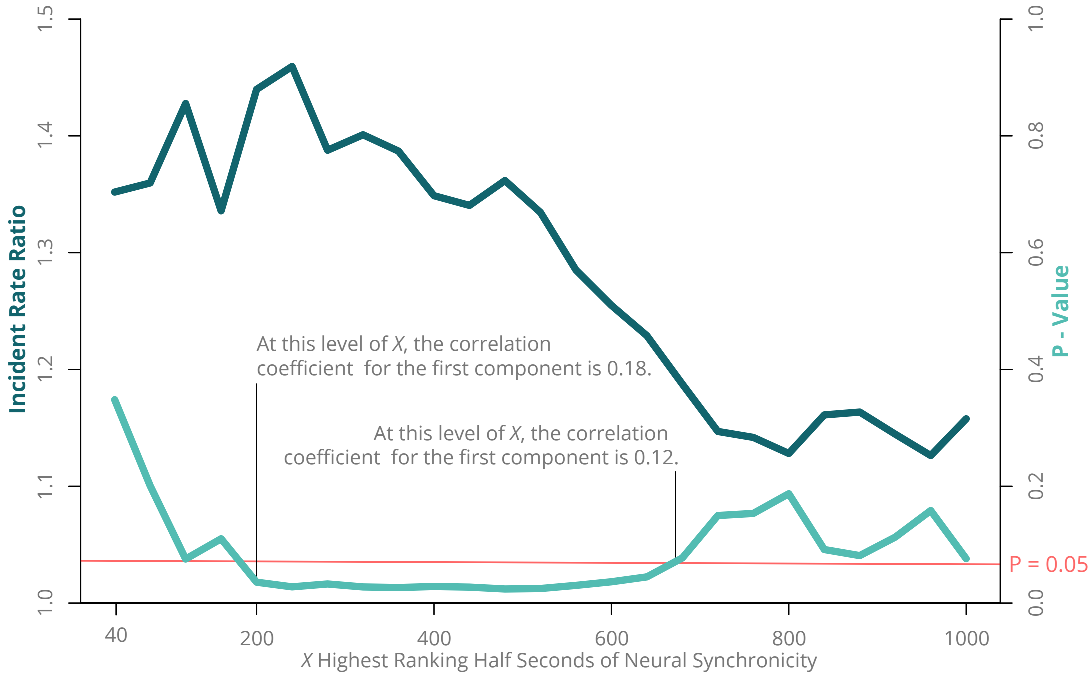
7 Results

7.1 Neural Engagement and Social Response

First, we assess the degree to which neural engagement predicts social resposne. We find that, in general, the first component of brain activity is a significant predictor of tweets, even when controlling for scene length and the presence of zombies and violence. The incident rate ratio (IRR) on the first component (Table 2, page 20) suggests that if a scene were to go from generating the lowest to the highest level of neural synchronicity, social media activity around that scene would increase by a factor of 1.57 while controlling for the scene's duration. This parameter drops to 1.41 when adding emotional variables (see Table 3), but remains significant at a level of $p<0.05$ across all models. While we include the second and third components in this model, they are marginally significant to insignficant, with the second component seeming to inversely vary with the outcome. This is a somewhat surprising result and future research is required before further interpretation. However, analysis of variance tests confirm that these two components do not significantly decrease the deviance in the model.

To further confirm these results, we also test how the effect size of neural engagement changes in moments with especially high correlations. Guided by past studies which indicate that spikes in correlated brain activity occur alongside espe-

Figure 7:
Incident Rate Ratios and P-Values from Bayesian Negative Binomial Regressions for Top X Ranked Windows of Neural Sychronicity on Weighted Tweets Per Scene



cially engaging content (outlined in Section 4 above), we should expect to see the effect size crescendo for scenes which feature extreme moments of synchronicity and diminish as these levels subside. In Figure 7 we visualize the results of an alternative to the model presented in Table 2 in which the EEG data is operationalized by assigning a "1" to a half-second window of brain activity if it falls within the top X windows across the entire the show and a "0" if not. Following our expectations, increasing values of X directly correspond with decreasing effect sizes. As the chart demonstrates, the IRR of the highly ranked windows of neural synchronicity slowly rises to a peak of 1.45 as X moves from 40 to 240 (or about the top 0.5% of observations to the top 2.5% of observations). This relationship concurrently increases in significance, with the p-value quickly approaching zero.

At this point, the cut-off level for the correlation coefficient of the first component is 0.18. The IRR continues to decrease as X moves from 240 to 680 (or about the top 2.5% of observations to the top 10% of observations), reaching a low of 1.15, while the p-value remains the same. At this point the effect dissappears and the relationship loses significance when the cut-off level for the first componenet drops below 0.12. These dynamics offer compelling evidence that spikes in neural synchronicity are correlated with social repsponse.

7.2 Emotion, Immersion, and Social Response

Examining the relationship between the overall emotional reaction to a scene and the level of associated social media activity (Tables 3-5, page 20), we find that, while holding all else constant, moments of the show that generated higher levels of intense, humorous, or personal reactions produced a significantly higher level of overall social media activity. Of these, personal involvement in the show's narrative elicits the strongest effect, and a move from a scene that generates no personal reactions to one which generates all personal reactions should increase the rate of tweets for that scene by a factor of 11.28. In Table 3 we also see that commentary on the show with regards to a particular scene is significantly and positively correlated with social media activity. Interestingly, in this model sentiment is inversely correlated with tweets. Here, the IRR of 0.27 on "sentiment" implies that scenes which generate solely positive reactions will lead to a decrease in overall social media activity by a factor of 3.7 (1/0.27).

We further explore the relationship between sentiment and commentary on the show by interacting these two variables in Table 4. In this model, the IRR's on

Results from Negative Binomial Regressions on Weighted Tweets

Table 2: Simple EEG Model

Coefficient	IRR	Estimate	Std. Error	T-Value	P-Value
(Intercept)	4.46	1.50	0.07	19.82	0.00
first component	**1.57**	**0.45**	**0.19**	**2.33**	**0.02**
second component	0.69	-0.37	0.18	-2.02	0.04
third component	1.12	0.11	0.20	0.55	0.58
violence	2.99	1.10	0.10	11.08	0.00
zombie(s)	3.71	1.31	0.05	25.85	0.00
scene duration	3.50	1.25	0.09	14.24	0.00

Table 3: Added Emotional Variables

Coefficient	IRR	Estimate	Std. Error	T-Value	P-Value
(Intercept)	2.04	0.71	0.06	12.40	0.00
first component	**1.41**	**0.35**	**0.14**	**2.40**	**0.02**
intensity	5.46	1.70	0.06	29.03	0.00
sentiment	0.27	-1.31	0.07	-19.85	0.00
humor	3.33	1.20	0.08	14.67	0.00
personal	12.03	2.49	0.06	40.38	0.00
show	1.29	0.26	0.05	4.91	0.00
zombie(s)	1.57	0.45	0.04	11.01	0.00
violence	4.64	1.53	0.07	21.50	0.00
scene duration	1.28	0.24	0.07	3.60	0.00

Table 4: Added Interaction

Coefficient	IRR	Estimate	Std. Error	T-Value	P-Value
(Intercept)	1.08	0.08	0.07	1.21	0.23
first component	**1.37**	**0.32**	**0.14**	**2.27**	**0.02**
intensity	8.71	2.16	0.06	34.50	0.00
sentiment	0.83	-0.18	0.09	-2.04	0.04
humor	2.34	0.85	0.08	10.41	0.00
personal	11.28	2.42	0.06	40.46	0.00
show	7.48	2.01	0.11	18.42	0.00
zombies(s)	1.57	0.45	0.04	11.37	0.00
violence	4.05	1.40	0.07	20.23	0.00
sentiment : show	0.02	-3.81	0.23	-16.85	0.00
scene duration	1.26	0.23	0.07	3.53	0.00

Table 5: Immersion Model

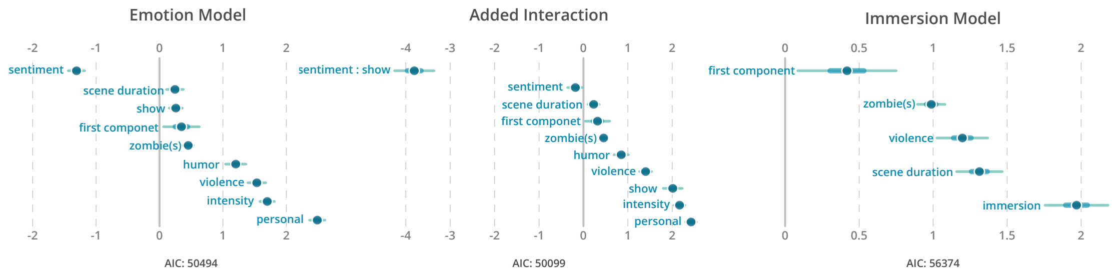
Coefficient	IRR	Estimate	Std. Error	T-Value	P-Value
(Intercept)	1.32	0.28	0.08	3.38	0.00
first component	**1.52**	**0.42**	**0.17**	**2.47**	**0.01**
immersion	**7.16**	**1.97**	**0.11**	**18.21**	**0.00**
violence	3.31	1.20	0.09	13.72	0.00
zombie(s)	2.68	0.99	0.05	20.99	0.00
scene duration	3.72	1.31	0.08	16.75	0.00

"sentiment" and "show" represent the factors by which social media will increase if one is high and the other is low. So the IRR of 0.83 on "sentiment" suggests that scenes which produce positive comments and do not reference the show increase the rate of tweets by a factor of 1.20 (1/0.83). Conversely, the IRR on "show" implies that the rate of tweets per scene will increase by a factor of 7.48 during scenes that generate negative comments about the show. Finally, the interaction term, "sentiment : show", shows the effect of scenes that elicit positive commentary about the show. The extremely low IRR of 0.02 suggests that such scenes decrease the rate of tweets by a factor of 50 (1/0.2) when compared to scenes which evoke negative commentary unrelated to the show. While we caution against over-interpreting this result – only 5% of tweets in the sample were coded for both positive sentiment and commentary on the show – we think this finding provides cautionary evidence against the use of positive sentiment as an indicator of social media engagement. Indeed, from these results, we would suspect that scenes which elicit negative commentary about the show's production actually lead to more tweets than those which evoke reactions of praise. We will expand upon these insights in Section 8 below.

Finally, in Table 5 we replace all emotional indicators with our metric for immersion. As hypothesized, this algorithmic combination of "intensity", "sentiment", "personal", "humor", and "show" is strongly and significantly correlated with social media activity, even when controlling for neural synchronicity and other relevant content variables. The IRR for "immersion" suggests that scenes which elicit highly immersed responses will increase tweets by a factor of 7.16 when compared to scenes that evoke wholly negative immersion.

However, when comparing this model with the previous two models, we find

Figure 8: Plots of Coefficients for Emotion and Immersion Models



that the level of variance accounted for by this model decreases when replacing all emotional indicators with immersion. Figure 8 visualizes the coefficients and confidence intervals for each of the three models with their associated Akaike Information Criteria (AIC) (in general, a lower AIC implies a better fitting model, however this is not always the case). The higher AIC for our immersion model suggests that our operationalization of immersion does not fully account for the variance in tweets per scene explained by our five indicators of emotion. To ensure that our operationalization of immersion is valid, we peform a Principal Components Analysis on the five emotional variables. This method attempts to find latent components hidden in multivariate data and estimate how much of the total variance is accounted for by each component. In this case we find that the first component of the emotional indicators accounts for about 55% of the total variance. This analysis offers cautious evidence for the existence of an underlying dimension of emotion. Furthermore, this component is loads heavily on intensity, the factor to which we assign the greatest weight in our calculation.

8 Discussion

Our results show preliminary evidence for a link between neural stimulus and social response. With further refinement and replication, content creators or producers could harness these methods to forecast what moments of their shows will elicit the most discussion on social media. Such analyses could inform the placement of advertisements and promotions. In time, researchers could use the methods outlined in this paper to isolate the neural signatures of various psychographic or demographic clusters and predict how these audience segements will respond to narrative moments differentially on social media.

A more immediate outcome of these results is an empirically-based critique of sentiment as the preferred metric for audience engagement on social media. There are certainly some industries - Air Travel, for instance - where sentiment is a useful indicator. But in the context of entertainment, the emotional valence of reactions seems to be meaningless without considering the particular context in which these emotions are expressed. As this paper finds, negative comments about a scene were actually one of the strongest predictors of overall social media activity. That users are even driven to make such comments suggests some level of engagement with the content. In turn, rather than relying on easily-implementable sentiment classificaiton algorithms, entertainment-focused social media analysts should strive to develop metrics that more accurately capture immersion.

Yet despite our positive results, there is no denying the relative weakness of the effect sizes. Most notably, the IRR for the first component of brain activity suggests that moving from the lowest to the highest levels of neural synchronicity only increases social media activity for a scene by a factor of 1.5. That means, all else equal, a scene that generates absolutely no neural correlation might see 10 tweets while a scene that generates the highest level of neural correlation in our sample would recieve 15 tweets. Furthermore, when we applied cross-validation techniques and attempted to train a model to predict social media activity at the level of a scene, we found that, while we were able to decently predict the relative rankings of scenes, our models were extremely poor at estimating the absolute per-scene rates of tweets.

Given these limitations, we turned to qualitative methods to further assess the exceptions to our model. In particular, we wondered what factors were associated with scenes that were neurally engaging but people didn't tweet about – a phenomenon we deem "ghost engagement." In Table 6, we see the top ten scenes ranked in terms of the number of windows of brain activity that fell in the top 3% of all windows over the log of total windows in that scene. The break-point of 3% was determined by identifying the point in Figure 7 where the effect size of the first component was the largest).

Table 6: Top Scenes by Neural Synchronicity

Spikes	/Log(N)	Tweets	Description
4	3.40	17.20	Shootout between Rick and suspects
4	2.90	54.40	Duane hits Rick in the head with a shovel
2	2.60	5.10	Rick is shot by suspects in back
4	2.50	5.30	Morgan threatens Rick with gun
4	2.50	12.70	Suspects' car rolls after hitting spike strip
3	2.20	49.50	Morgan shoots zombie in the head
3	2.10	1.70	Rick is shot by suspects in body armor
3	1.90	1.10	Morgan reassures Duane while shooting zombies
4	1.90	69.60	Rick shoots half- zombie in the head
2	1.80	132.10	Rick shoots little zombie girl in the head

As Table 6 indicates, many of the top neurally engaging scenes also generated many tweets. However, a major exception was a series of scenes involving a car chase and shootout towards the beginning of the show. While these scenes produced many spikes in brain activity, they saw relatively few tweets. We suspect this may have something to do with the fact that these moments did not involve zombies – a novel element of the show that users were most likely predisposed to tweeting about. On the flip side, a series of scenes towards the end of the show involving a horse being eaten by zombies recieved nearly 1/3rd of all tweets in our sample yet produced no spikes in neural synchronicity. This disconnect between neural engagement and social response is most likely because of the jarring nature of the scene. Furthermore, this scene generated an overwhelmingly amount of negative sentiment as a large number of people were horrified by the depiction of an innocent horse being eaten by a horde of zombies.

Finally, a couple of scenes involving highly emotional moments saw almost no tweets but a high level of neural synchronicity. These scenes involved the main character sobbing when he suspects his family is dead and a supporting character breaking down in tears when he sees his mother as a zombie, respectively. These anecdotes seem to suggest support for Berger and Milkman's hypothesis that emotionally deactivating content discourages sharing. However, without an objective schema for the emotional characteristics of each scene, we cannot yet make any definitive conclusions.

9 Conclusion

Through its novel combination of neuroscience with content and social media analysis, this study offers a unique perspective on the question of what constitutes audience engagement. In particular, it outlines an innovative methodology that enables the rigorous comparison of narrative elements, neural synchronicity, and social response throughout the course of a television show. Through the application of this methodology to the series premiere of *The Walking Dead*, this study finds that neural synchronicity is significantly correlated with social response. This relationship appears to be especially strong in moments when the audience's neural signals spike concurrently. The correlation of these two indicators suggest that their combination may lead to more meaningful metrics of audience engagement.

Looking at the emotions elicited by the show, our models suggest that scenes which evoke intense personal and/or humorous reactions to content are strongly associated with more activity on Twitter, even when controlling for neural synchronicity and relevant content variables. Interestingly, scenes that generate negative commentary about the show are far more likely to generate social media activity than those which evoke positive comments about the show. This suggests that the use of sentiment as an indicator for audience engagement is potentially unfounded, as it ignores the context in which these emotions are expressed. Finally, by combining emotional indicators into an index of immersion that weights intense comments expressing personal investment in the narrative over matter-of-fact commentary on the show, we find that immersion is a strong predictor of social response. We hope this finding opens a path for the development of better schemas for classifying emotions embedded in social media messages.

However, while our models suggest a link between neural synchronicity and social response, the effect size is relatively weak. Further investigation through a visual representation of our data sources reveals anecdotal evidence for the presence of "ghost engagement", or moments of the show which are neurally stimulating but do not generate much activity on Twitter (or vice versa). While we speculate that these examples are explained by the emotional salience or novelty of the content in these scenes and/or their temporal placement, further research is required before we can make any definitive conclusions. Here, the development of methodologies and taxonomies for rigorously classifying the emotions evoked by a narrative will be particularly useful.

References

[1] Berger, J. A. and Milkman, K. L. (2009). What Makes Online Content Viral? Available at SSRN: `http://ssrn.com/abstract=1528077`.

[2] Dowling, E. (2012, October 17). Second Presidential Debate Tops Super Bowl in Social TV. Message posted to `http://bluefinlabs.com/blog`.

[3] Darabont, F and Kirkman, R. (Writers), and Darabont, F. (Director). (2010). Days Gone By. In The Walking Dead. New York: AMC.

[4] Dmochowski, J. P., Sajda P. , Dias J. , and Parra, L. C. (2012). Correlated components of ongoing EEG point to emotionally-laden attention: a possible marker of engagement? Frontiers in Human Neuroscience, 6:112.

[5] Greenmeier, L. (2013, January 3). Nielsen and Twitter Team to Track TV. Message posted to `http://www.scientificamerican.com`.

[6] Hasson, U., Nir, T., Levy, I., Fuhrmann, G., Malach, R. (2004). Intersubject Synchronization of Cortical Activity During Natural Vision. Science, 303:5664.

[7] Jannrone, J. (2012, September 17). When Twitter Fans Steer TV. Message posted to http://online.wsj.com/.

[8] O'Brien, T. (2010, April 3). Scientists Predict Box Office Revenue With Twitter. Message posted to http://www.switched.com/.

[9] Russell, J. (1980). A Circumplex Model of Affect. Journal of Personality and Social Psychology, 39:6.

[10] Subramanyam, R. (2011, October 6). The Relationship Between Social Media Buzz and TV Ratings. Message posted to http://blog.nielsen.com/.

[11] Twitter. (2012). Twitter on TV: A Producer's Guide. Retrieved December 12, 2012, from https://dev.twitter.com/media/twitter-tv.

[12] Whitaker, Z. (2012, July 30). Twitter: Olympics opening ceremony saw 9.66 million mentions. Message posted to http://news.cnet.com/.