Trailer Brain: Neural and Behavioral Analysis of Social Issue Documentary Viewing with Low-Density EEG*

Jason S. Sherwin[1], Corinne Brenner[2], and John S. Johnson[3]

1 **Harmony Institute, New York, NY, USA**
jason@harmony-institute.org
2 **Harmony Institute, New York, NY, USA**
3 **Harmony Institute, New York, NY, USA**

Abstract

The effects of social issue documentaries are diverse. In particular, monetary donations and advocacy on social media are behavioral effects with public consequences. Conversely, information-seeking about an issue is potentially done in private. We designed a combined free-viewing and rapid perceptual decision-making experiment to simulate a real scenario confronted by otherwise uninformed movie-viewers, i.e., to determine what degree of support they will lend to a film based on its trailer. For a cohort of subjects with active video-streaming (e.g., Netflix) and social media accounts (e.g., Facebook), we recorded electroencephalography (EEG) and behavioral responses to trailers of social issue documentaries. We examined EEG using reliable component analysis (RCA), finding reliability within subjects across multiple viewings and across subjects within a given viewing of the same trailer. We found this reliability both over EEG captured from whole-movie viewing, as well as over 5-second movie segments. Behavioral responses following trailer viewing were not consistent from first to second viewings. Rather, support choices both tended towards extremes of support/non-support and were made faster upon second viewing. We hypothesized a relationship between reliability behavioral metrics, finding credible evidence for it in this dataset. Finally, we found that we could suitably train a naive classifier to categorize production value and narrative voice ratings given to the viewed movies from RCA-based metrics alone. In sum, our results show that EEG components during free-viewing of social issue documentary trailers can provide a useful tool to investigate viewers' neural responses during viewing, when coupled with a post hoc behavioral decision-making paradigm. The possibility of this tool being used by producers and filmmakers is also discussed.

1998 ACM Subject Classification F.1.1 Models of Computation, I.6 Simulation and Modeling, J.3 Life and Medical Sciences

Keywords and phrases EEG, reliable components analysis, machine learning, documentary films

Digital Object Identifier 10.4230/OASIcs.CMN.2016.2

1 Introduction

Video media have the potential to manifest an almost infinite variety of effects on society, over a long timescale. This is because the conduits of such effects are human viewers, each bringing a different nervous system to bear on the sensory stimulus of a given video. Furthermore, both the sensorial complexity of visual-auditory stimuli and the culturally-laden content in

* This work was supported by the Rita Allen Foundation and the Harmony Institute.

7th Workshop on Computational Models of Narrative (CMN 2016).
Editors: Ben Miller, Antonio Lieto, Rémi Ronfard, Stephen G. Ware, and Mark A. Finlayson; Article No. 2; pp. 2:1–2:21
Open Access Series in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

such videos (e.g., visual and auditory semantics) make the connection between stimulus and nervous system response a challenging relationship to capture.

Social issue documentary trailers provide a useful stimulus set to study these effects because they are created to elicit action on the part of the viewer. Representative actions covering a range of economic, social and educational effects, can be quantified. Large user-bases of online video-streaming services (e.g., Netflix) allow instant access to more information on a given topic via viewing the full movie advertised in the trailer. Similarly, the large user-bases of social media networks (e.g., Facebook) allow support (e.g., a Facebook 'Like') of any shared content. Due to the currency market functionally being an even larger user-base than the previous two, and the frequency with which social issue documentaries are associated with requests for donations to specific actions or causes, monetary spending behavior becomes yet another measure to assess a trailer's effects. Consequently, the effects of social issue documentary trailers on a large population can be assessed on a subject group pulled from these overlapping user-bases.

Many studies have investigated potential relationships between behavioral and neural response to video media [9, 10, 12, 13]. For example, Kato et al. claimed changes in functional magnetic resonance imaging (fMRI) response of prefrontal and medial prefrontal cortices among subjects supporting a political candidate when viewing negative video advertisements about that candidate [18]. Aside from methodological concerns about this study (e.g., uncorrected p-values for fMRI blood oxygen-level dependent (BOLD) maps), this study considers how video media affects neural response when the decision to support a given candidate has already been made. In other words, there is no behavioral consequence of the video in this case. In another study of videos about political candidates, Zhang et al. found that neural response and computer vision analysis that revealed rhetorical gestures of the candidates were related [34]. However, as with Kato et al., this study recruited subjects on the basis of their prior support of the political party of either candidate (e.g., Democrat or Republic for the 2012 U.S. Presidential Campaign). Furthermore, there were no behavioral responses that followed movie-viewing and, even if there had been, such responses would be irrelevant because the election had already taken place.

The techniques of neuroimaging analysis for movie-viewing are also varied. Hasson et al. has shown that reliable component analysis (RCA) applied to fMRI is a viable tool for analyzing such data [12], finding relationships between RCA-based metrics and viewer/listener comprehension [11, 30], as well as time-resolved memory-encoding measured with both electrocorticography (ECoG) and fMRI [6, 10]. While some of these studies recorded behavioral responses (e.g., in the form of post hoc comprehension surveys [9, 13]), there was no element of support either for or against the viewed content embedded in these responses. In a variation of RCA that includes dimensionality reduction as a preliminary preprocessing step to recorded electroencephalography (EEG), Dmochowski et al. found that crafted video narratives (e.g., scenes from Hollywood feature films) elicited more reliable component activity in EEG than control videos [5]. In follow up work to this result, Dmochowski et al. found that such activity in a subject cohort (i.e., a small group) could be related to mass social media response for viewers (i.e., a large group) who naively watched the same video for the first time [3]. Both the studies of Hasson et al. and Dmochowski et al. occurred within precisely controlled and otherwise uncommon viewing environments (e.g., laying down in an fMRI machine or enclosed within a room shielded for radio frequency interference).

While there is no direct mapping yet between candidate narrative theories and neuroimaging analysis, RCA and related techniques provide a possible method for investigation. But RCA-based techniques frame the problem of narrative characteristics and content delivery on the sensor and/or voxel level, looking at the consistently reliable activity in that measurement. As Honey et al. and others have shown [14, 28], the flexible length of temporally receptive windows (TRWs) in the cortex mediates a hierarchical comprehension of complex temporal stimuli (e.g., movies, or video narratives). The extent to which these hierarchies of comprehension map to simple narrative content choices, such as narrative voice (e.g., participatory or expository) and production style (e.g., as manifested by high, medium or low budget productions), has yet to be uncovered.

Other work has shown that content characteristics manifest in audience comprehension of the viewed media and subsequent choices – explicitly or implicitly made – concerning that media. While RCA analyses have incorporated surveys of qualitative comprehension [12, 13, 10] and preference [3] post hoc, the impact of comprehension on decision time has not been investigated. Though largely ignored in previous work, the focus on decision time is important because it can be an indication of cognitive conflict when posed with alternative choices [19].

There has been some academic work on the topic of audience response to movie trailers, but not with the same rigor as seen for other video media. For example, Jerrick utilized a survey approach to gauge film trailer effectiveness in the U.S. college student market [17]. In another study, Findsterwalder et al. used post hoc interview transcripts as the basis for a qualitative analysis of movie trailer effectiveness in the New Zealand cinema market [8]. Neither of these studies examined audience response at the fine level of detail allowed by biometrics, such as EEG or fMRI. The Jerrick study did not analyze the effects of the movie trailer viewing on follow-up behavior with regards to having seen that video, while the Findsterwalder et al. study only did so through a qualitative analysis of interview transcripts.

In this paper, we describe how an experimental paradigm combining free-viewing of movie trailers and a follow-up decision-making task allows us to analyze the behavioral and neural responses elicited by an otherwise naive viewing. To do so in a more natural viewing environment, we use a wireless low-spatial density EEG outside of a radio-frequency-shielded environment to measure neural response during movie trailer viewing. By following each trailer viewing with an alternative forced choice (AFC) task whose consequences are linked to personal video-streaming and social media accounts, as well as a fixed monetary endowment, we gauge an immediate level of support for the viewed trailer that covers educational, social, and economic support categories of interest, respectively. We use proven techniques of RCA, as applied to EEG by Dmochowski et al. [5, 3, 17], to track reliable activity across both multiple viewings and/or subjects on both whole-movie and sub-movie (e.g., 1 second) time scales. Finally, we use regression and other machine learning techniques to link behavioral, neural and/or narrative features of the viewed movies.

2 Methods

2.1 Subjects

12 subjects were recruited for this study and data from 2 subjects were unusable due to EEG hardware malfunction. Of the 10 remaining subjects (5 male), the age range was 32.6 ± 2.0 years. Informed consent was obtained from each of them in accordance with the guidelines and approval of the IRB Solutions Institutional Review Board. Each subject was told they were to be paid $20 for their participation in this study. The requirements for participation were that each subject had at the time of the experiment an active movie-streaming account (e.g., Netflix) and a Facebook account to which he/she must log in at the experiment's start. Also, no subject should have viewed two or more of these movies in their entirety before recruitment to this experiment.

Table 1 Title, production value (PV) and narrative voice (NV) of all movies' trailers.

Movie No.	Title	Theor, PV	Alg. PV	Theor. NV	Alg. NV
1	A River Changes Course	high	mid-high	expository	expository
2	Cool It	mid	mid-high	participatory	participatory
3	Bag It	mid	mid-low	participatory	participatory
4	No Impact Man	low	low	participatory	part.-obs.
5	YERT	low	low	participatory	participatory
6	Carbon Nation	mid	mid-high	expository	expository
7	Fall and Winter	high	high	expository	expository
8	Fight for the Planet	low	low	expository	exp.-part.
9	Home	high	mid-high	expository	expository

2.2 Stimuli Overview

We presented movie trailers using Matlab and PsychToolbox [1] on a Dell Inspiron 15 (5000) 4th-Gen Core i7 laptop. We selected movie trailers from a database of over 435 films that constitutes the StoryPilot database [31]. The StoryPilot database catalogues social issue documentary films across a broad range of topics, noting such film characteristics as social issue topic, production value (PV), narrative voice (NV) and other metrics.

Researchers independently classified all films in this database for social issue content, PV and NV, among other metrics, based on predetermined criteria. NV values were based on a classification scheme developed by Nichols [24]. Of the entire database, inter-coder reliability was assessed for 79 films of the total sample coded by 2 or more researchers. Inter-coder reliability statistics for PV were 83.1% agreement, Cohen's kappa: 0.70; For NV, coding was resolved by discussion, so agreement was 100%.

From this sample, we selected nine (9) films on the social issue topic of the environment. This was done to control for any audience priority differences across different social issues. The other decision criteria for selecting these 9 films were to cover a range of low, mid, and high PVs, as well as both expository and participatory NVs, as logged in the StoryPilot database (i.e., the theoretic values for PV and NV in Table 1). The movies whose trailers were selected for viewing are listed in Table 1.

Table 1 also shows algorithmic values for PV and NV. These values were determined from re-running the inter-coder reliability analysis on the original coding values that had gone into previous discussion-based analysis. We ran this calculation to remove any potential subjective bias resulting from group discussion of the selected movies' NV and PV values. Across two coders for NV, we found Cohen's kappa of 0.63. For PV, we calculated Cohen's across three coders both with and without allowance for variation by a value of one (i.e., low to mid, or high to mid). Such an allowance is suitable for PV because it is a monotonic value scale, unlike NV. Without allowance, we found Cohen's kappa values were 1.0, -0.03, and -0.03 among the three unique coder pairings. With allowance, we found Cohen's kappa values of 1.0, 0.62, and 0.62. Therefore, inter-coder reliability on NV was quite high among our selected film trailers, while it was strong amongst PV, though there was some variability. Based on these inter-coder results, the nine selected films covered our criteria of mono-topical films that covered a broad range of NV and PV values.

After each viewing, each subject was presented with a prompt screen displaying 4 possible behavioral responses to the movie trailer.

Table 2 Behavioral response choices and linked icons.

Action	Linked Icon
Donate $1 to a cause related to the viewed trailer	
‘Like’ this movie on their Facebook page	
Add the full movie to their movie-streaming queue	NETFLIX
Indicate ‘No Interest’ in the trailer topic	N/A

2.3 Behavioral Paradigm

Subjects were shown a sequence of nine (9) unique movie trailers at an inter-stimulus interval (ISI) determined by the subjects’ readiness. Subjects sat at a comfortable distance from the screen. The paradigm and data acquisition are illustrated in Figure 1.

In total, subjects viewed unique pseudo-randomized orderings of the trailer sequence twice. In particular, all trailer viewings were randomized, so that no subject would see the exact same ordering of movie trailers.

The prompt screen contained visual icons, each representing different possible ways of supporting or not supporting the previously viewed movie trailer (middle frame of Figure 1A). After each movie trailer, the subject was instructed to select from four possible actions in response to the previously viewed trailer as shown in Table 2.

Subjects logged into Facebook and movie-streaming accounts at the start of the experimental session to simulate actual consequences of support choices. Subjects were also alerted to this forthcoming choice, as well as to which keyboard buttons must be used to indicate his/her choice, at the start of the experimental session. Each subject was told to respond as quickly as possible. Choice-button relationships were pseudo-randomized for each choice screen to counter a possible habituation effect in choice selection.

2.4 Data Acquisition and Preprocessing

EEG data was acquired without electrostatic shielding using an Advanced Brain Monitoring X10 9-channel system (Carlsbad, California) with scalp electrodes arranged in the 10-20 System. Data was sampled at 256 Hz. A software-based 0.5 Hz high pass filter was used to remove DC drifts and a 30Hz low pass filter was used to isolate relevant EEG power bands. These filters were linear-phase to minimize delay distortions. Stimulus events, i.e., movie start, movie frame flips, movie end, choice screen and keyboard responses, were recorded on separate channels.

Two preprocessing steps were employed on the EEG data. First, each channel was z-scored. Second, to ensure that all subjects were on the same timeline for each movie, the movie frame

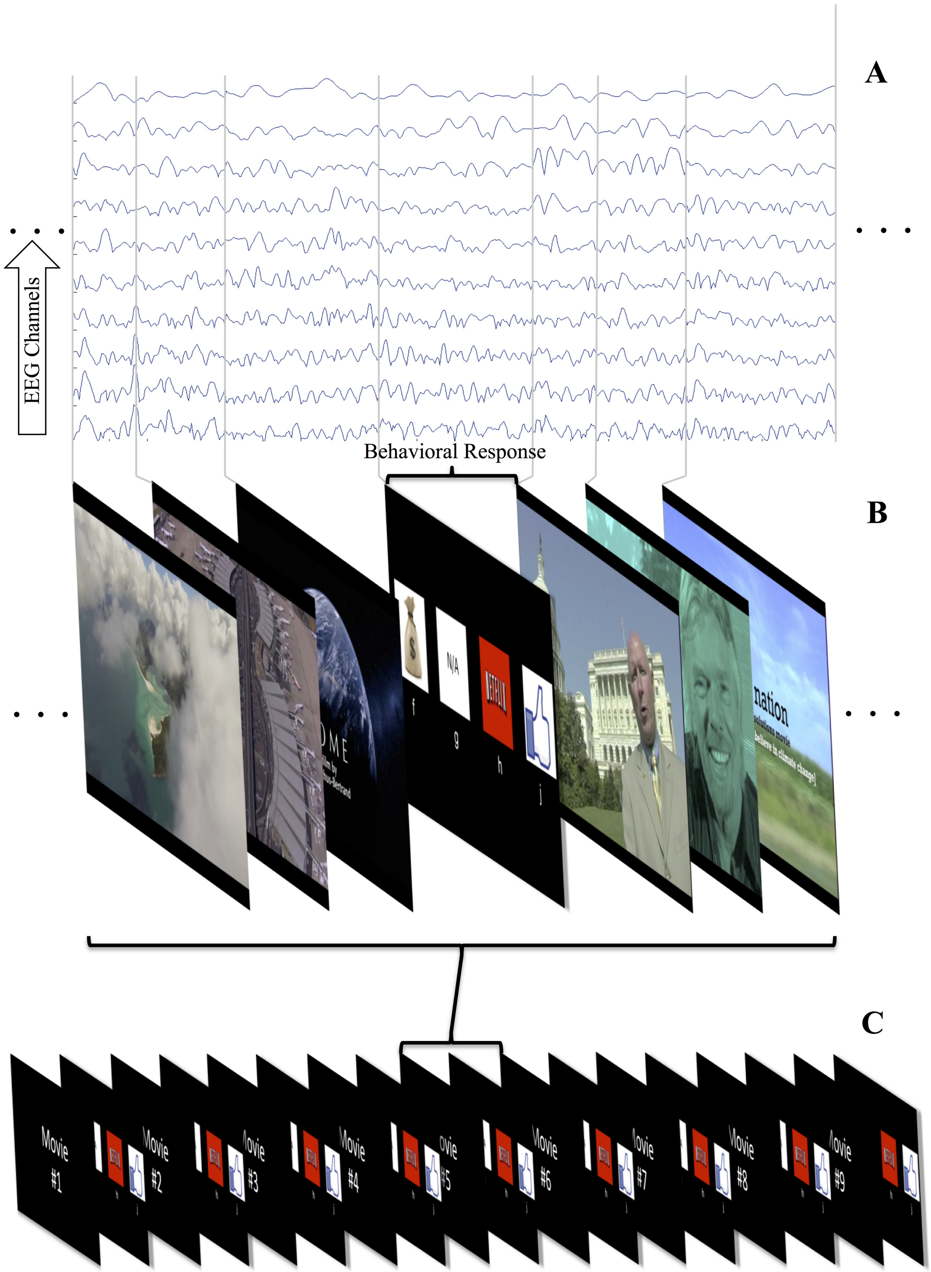


Figure 1 Paradigm overview. Frames of chosen movie trailers are shown (B) with events marked across all EEG channels (A). Horizontal ellipses before and after movie frames indicate previous and subsequent movie frames (B) across EEG time series (A). Each row of signal represents a different EEG channel from the n = 9 channels used here. Following each movie viewing, there was a prompt message for the subject to get ready and proceed (not shown) to the choice screen (middle frame of B). Response time was recorded from presentation of the choice screen as indicated. The subsequent movie then began playing. A total of nine (9) movies, each with prompt/response screens comprised one full block (C). Two blocks were run per subject in succession. Movie sequences were pseudo-random and so was choice key matching at each response screen (e.g., 'add to Netflix queue' could be in either of the four positions shown after each viewing).

flip events were used as a marker between each of which the EEG data was time-averaged. This resulted in one EEG measurement per frame flip. This method of time-locking EEG to movie stimuli counteracts slight potential variability in EEG hardware sampling rates that can cause substantial drift as recording enters multiple minutes. Also, the variability in the refresh rate of any PC video card is mitigated with this technique. We considered the use of Independent Components Analysis (via the 'fast ica' algorithm) to remove eye-blink and eye-movement artifacts, but saw no noticeable differences on the EEG data when this technique was inserted before z-scoring.

2.5 Behavioral Analysis

Choice values and response times (RTs) from prompt were analyzed for each movie. As a starting hypothesis, choice values were assigned numerical equivalents as listed previously to gauge choice variability upon repeat viewing. Additional analysis employed a coding scheme, whereby choice-values of 1 and 4, i.e., 'Donate $1' and indicate 'No Interest', were grouped together, while choice-values of 2 and 3, i.e., ' 'Like' this movie' and 'Add the full movie to queue', were also grouped together. Both this binary coding scheme and response times were analyzed with standard t-tests to gauge differences from first- to second-viewing.

2.6 EEG Analysis Overview

Our primary method for analyzing EEG was reliable component analysis (RCA) [5, 3, 4]. RCA is a technique used to analyze neural data acquired during presentation of continuous stimuli, for which there is no canonical locking event (e.g., as in standard P300 or oddball EEG paradigms). In general, RCA requires at least two neural data signals $X_1 \in \mathbb{R}^{DxT}$ and $X_2 \in \mathbb{R}^{DxT}$, where D is the number of channels and T is the number of time samples. RCA seeks a weight vector ($\mathbf{w}$) such that the resulting linear projections $\mathbf{y}_1 = X_1^T\mathbf{w}$ and $\mathbf{y}_2 = X_2^T\mathbf{w}$ have maximal correlation. Following the technique of Dmochowski et al. [4], we selected the weight vectors that yielded the three highest correlations after linear projection. These projections are the three most reliable components, henceforth to be known as 'the reliable components'. We applied this technique to different subsets of the overall neural data, both for whole movie-viewing and sub-clips.

2.7 EEG Analysis: RCA with Whole-Movie Viewing

To examine common neural activity between repeat viewings, we assumed X_1 and X_2 to be first- and second-viewings of the same movie trailer. In this manner, we calculated the reliable components for each subject viewing each movie trailer. We then validated the significance of these components by bootstrapping. In particular, we time-shuffled one of the signals before recalculating the reliable components, keeping the correlation value from each shuffle. We performed N = 10,000 such calculations, for each subject and each movie, so that the true correlation value determined from RCA could be compared to the distribution of values obtained from the time-shuffle-based calculation. A reliable component was deemed significant if the correlation value it yielded between X_1 and X_2 exceeded p = 0.05 (two-tailed, p = 0.025). Both negative and positive correlations are possible.

To examine common neural activity across subjects within the same viewing, a similar calculation was made, except all subjects' data signals were concatenated such that each subject's data was paired with every other subject's data for a given viewing (see Table 3). Color-coding in the table indicates telescoped blocks by which all subject combinations were formed to make X_1^{all} and X_2^{all}.

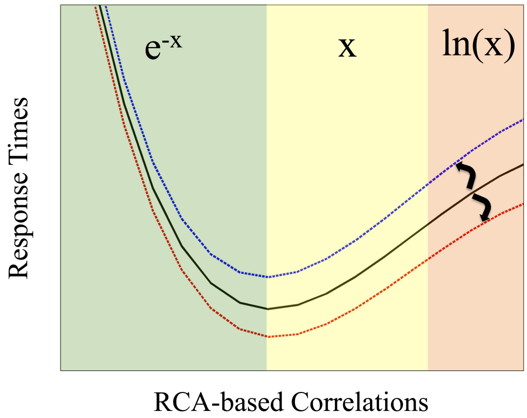


Figure 2 Hypothesized relationship between RCA-based correlations and response times. Three portions of variability of response times are shown in green, yellow and orange shading. The exponential decay portion (green) represents high response times when correlations are low, while the logarithmic portion (orange) represents mid-range response times when correlations are high. Lowest response times are hypothesized for the mid-range correlations connected to the extremes by a linear relationship (yellow) between dependent and independent variables. Variation of the constants multiplied by these three portions will produce variability of the curve as exhibited by the red and blue dashed curves around the central black curve.

Table 3 Illustration of concatenation for all-subject RCA.

$X_1^{all} \equiv$	1	1	...	1	2	2	...	2	...	8	8	9
$X_2^{all} \equiv$	2	3	...	10	3	4	...	10	...	9	10	10

The reliable components from the newly formed X_1^{all} and X_2^{all} were then calculated. As for individual subjects, the bootstrapping provided a threshold to evaluate each reliable component's significance.

2.8 EEG Analysis: RCA with Time-Resolved Viewing

To examine common activity between multiple viewings in a time-resolved manner, RCA was also applied to sub-clips of all trailers. In particular, sub-clips of X_1 (first viewing) and X_2 (second viewing) were taken from 150-frame ($\approx$5s at screen refresh rate) time intervals and subject to RCA calculation, including bootstrapping for significance. This 150-frame time interval began at trailer onset and then was moved forward in time by 30 frames ($\approx$1s) to do another RCA calculation with bootstrapping. This procedure was repeated until the whole of X_1 and X_2 were covered. A correction by false discovery rate (FDR) was applied to each interval's p-value, due to the repetition of this procedure for the number of overlapping 150-frame windows in each movie trailer. The result of this calculation is a time-resolved map of reliable components, each with a significance value against a computed null distribution, for each movie trailer and each subject.

2.9 Relating EEG to Behavioral Responses

We examined a possible link between EEG and behavioral metrics. In particular, we tested the hypothesis that response times vs. RCA-based correlations followed a trend illustrated in Figure 2.

We hypothesized this relationship as a corollary to previous work linking higher correlation values to greater comprehension of a viewed movie [10, 13]. Particularly, in the context of a decision to choose among support levels for a viewed documentary trailer, we reasoned that the corollary of this result is that extremely low and high correlation values would lead to delayed reaction times, while mid-range correlation values would lead to the quickest reaction times. Because correlation values have been linked monotonically to movie comprehension, we reasoned that the low correlations would produce generally higher RTs than high correlations, thus the shape of the hypothesized relationship in Figure 2.

2.10 Classifying Narrative Voice and Production Value from EEG

We also examined a link between the EEG data and the narrative features of each movie trailer. In particular, we trained a decision tree classifier [29] on the results of the RCA from whole-movie viewing. We used the results of whole-movie viewing because these metrics were calculated from both all time points of the trailers and all subjects who viewed the movie trailers. This approach allowed us to sample as many time points and viewers as possible to judge narrative content of a given movie.

3 Results

3.1 Behavioral Choices More Extreme and Faster Upon Second-Viewing

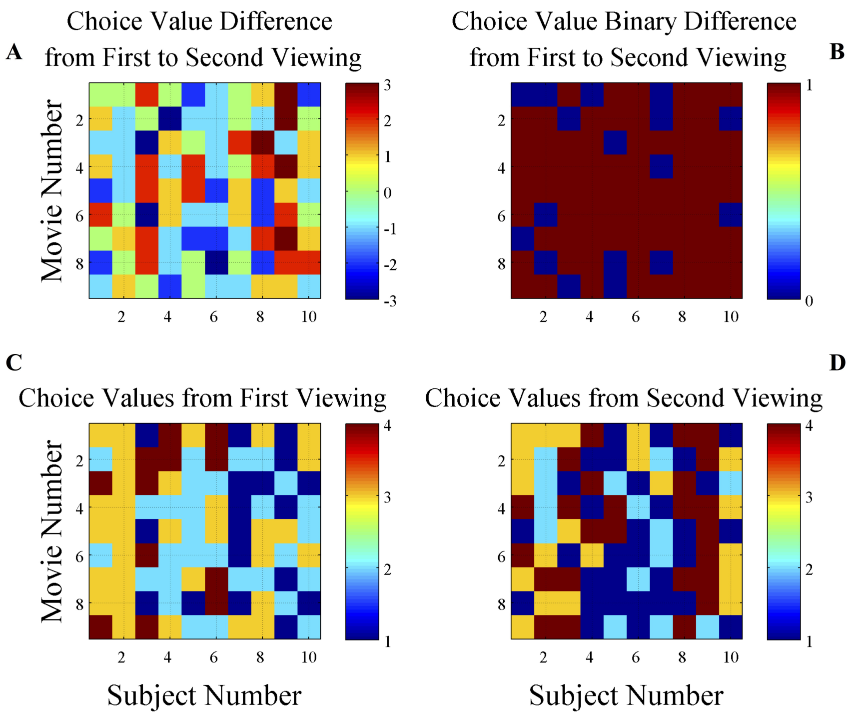
An overview of choice values for each subject and movie is shown in Figure 3. Not only does Figure 3 show a broad range of choice-values across most subjects and movies within each viewing (Figure 3C and Figure 3D), but it also shows the variability in the magnitude by which choice values fluctuate between first and second viewings (Figure 3A). Noting changed and unchanged choice-values with a binary coding scheme, we see in Figure 3B that the majority of behavioral responses were different from first to second viewing. The number of choice-value differences was different across movies (t-test, $p < 0.01$) and subjects (t-test, $p < 0.01$).

Not only did we find that choice-values changed from first to second viewing, but they also followed the trend shown in Figure 4A. In particular, choice-values on the whole went from mid-level support (e.g., 2's and 3's) to extreme-levels of support (e.g., 1's and 4's). We quantified this by grouping mid-level and extreme-level choice-values. We compared the number of mid- and extreme-level choices within each viewing, finding significant differences for each (first-viewing: t-test, $p < 0.01$; second-viewing: t-test, $p < 0.02$). We also compared the number of mid- and extreme-level choices across viewings, finding significant differences for each (mid-level: t-test, p « 0.01; extreme-level: t-test, p « 0.01).

We also examined response times once subjects were prompted for their choice-values. Figure 4B shows an overview of mean and standard error response times (RTs) for each movie. As the figure shows, first-viewing RTs were significantly greater than those of the second-viewing. We quantified this by grouping all RTs within each viewing and then comparing to those of the second viewing (first-viewing > second-viewing: t-test, $p < 0.02$).

3.2 EEG Analysis: Whole-Movie Viewing Shows Certain Films Having High RCA Correlation

We performed whole-movie RCA for each movie trailer and subject. Figure 5 shows correlation magnitudes' mean and standard errors across subjects within each movie.



■ **Figure 3** Overview of choice values by movie and subject number. The difference in choice value from first to second viewing (A) is shown with its own color scale covering [-3,3]. For instance, a choice value of 1 in the first viewing and a choice value of 3 in the second viewing would be a difference in choice value of 2 (e.g., Movie Number = 1, Subject Number = 3). The binary version of this plot is shown next to it (B) with its own color scale, indicating 1 when the choice value did not change and 0 when it did. The bottom row of matrices shows the actual choice values from first viewing (C) and second viewing, each with the same color scale covering [1,4]. Movie numbers are taken from Table 1.

From Figure 5, we see variability in each of the components' mean correlation magnitudes, while standard errors across movies within each component are consistent. Whole-movie RCA within each viewing also showed variability by movie trailer. Figure 6 shows an overview of significant (green) and insignificant (red) component/viewing pairings.

From Figure 6, we also see the variability of significant components for each movie. One movie ("Bag It") maintained significant correlations across subjects within both viewings for all components, while four movies had only one component insignificantly correlated across both viewing. All other movies had at least one component showing insignificant correlation in at least two component-viewing pairs.

3.3 EEG Analysis: RCA with Time-Resolved Viewing Shows Frame-by-Frame Audience Impact

We also tracked RCA correlation values at a time-resolution of 1s (Figure 7). Figure 7 shows the time course of significant windows for three components of both positive and negative correlations for a given movie and subject.

From Figure 7, we see that positive correlations (blue) dominate, but that negative ones (red) exist as well. We summarized these results across all subjects and movies by calculating the movie-length-normalized number of significantly correlated windows (Figure 8). Figure 8

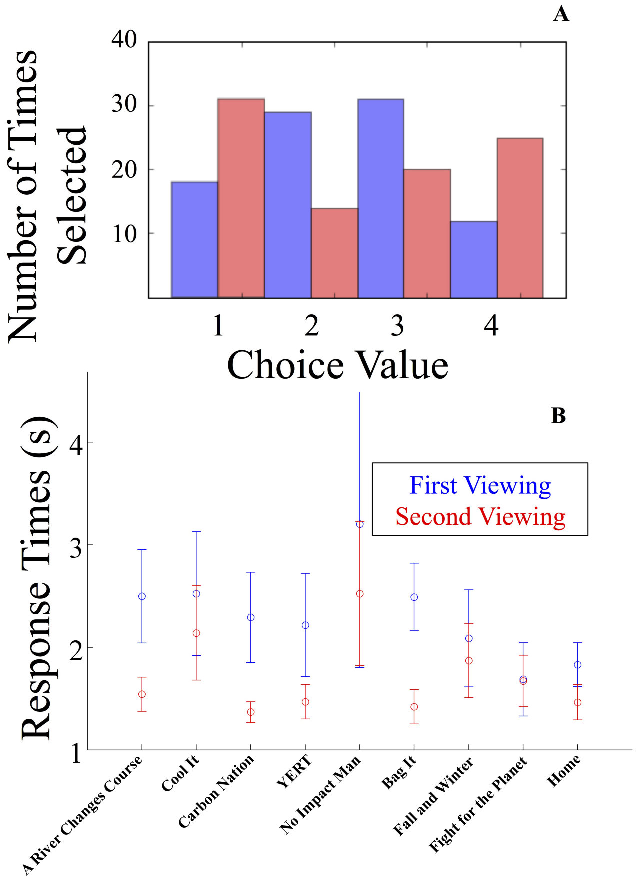


Figure 4 Overview of choice values across all movies and response times by movie. Viewing number is color-coded in the inset. Breakdown of choice values across all movies for first and second viewings is shown with the same color-coding (A). Mean response times and standard errors (B) are shown for first and second viewings. Error bars indicate standard error around the mean response time for the indicated movie. Across all movies, response times dropped from first to second viewing (t-test, $p < 0.02$).

shows mean and standard error number of significantly correlated windows for each movie, where the bar is color-coded to the movie trailer name. Positive and negative correlations are shown.

From Figure 8, we see that positive correlations are more frequent than the negative ones. But non-zero negative correlations were seen across all components, most notably from the first and third components, so they have been reported here, despite earlier work ignoring such results [5, 3].

3.4 RCA of EEG Relates to Behavioral Responses

We examined the relationship between EEG metrics obtained with RCA and behavioral responses (Figure 9). In particular, we tested the hypothesis that RCA correlation magnitudes would produce a relationship against response times (RTs) of the form, $y=Ae^{-x}+B\ln(x)+Cx+D$. First and second-viewings are temporally dependent events, so we examined each scenario alone, also considering the difference in RTs from first- to second-viewing.

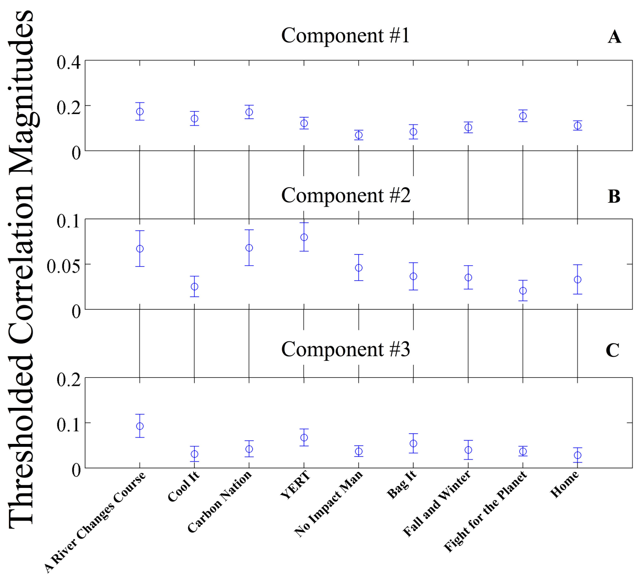


Figure 5 Correlation magnitudes across subjects for each movie. First (A), second (B) and third (C) components' mean correlation magnitudes are shown with error bars indicating standard error. Movie labels for the bottom plot (C) apply to corresponding plots above (A and B). Correlation values that were insignificant for a given subject-movie pair were zeroed out. Since magnitudes of significant correlation values are used, both positive and negative correlations from first- to second-viewing are represented here.

Figure 9A shows variation of first-viewing mean RTs with summed significant correlation magnitudes. High RTs are found for low correlation magnitudes. Low RTs are found for mid-range correlation magnitudes. And mid-range RTs are found for high correlation magnitudes. A similar trend is seen in Figure 9C, but here the mean difference in RTs is plotted. Also, differences in RTs are seen for both low and high correlation magnitudes. Both Figure 9A and Figure 9C show significant fits for the hypothesized variation of the behavioral with the neural metric. Figure 9B shows a trend in which second-viewing RTs fall off nearly monotonically with correlation magnitudes. At this sample size though, the fit is not significant.

3.5 Narrative Features of Trailers Classified from EEG

We also examined the relationship between EEG metrics obtained with RCA and the narrative features of the movie trailers, such as Production Value (PV) and Narrative Voice (NV), as shown in Table 1. As labels, we used the clear categories defined by theoretic NVs and PVs. We did this in order to have clear training labels that did not straddle multiple categories. The RCA values we used made a 6-component feature vector for each movie, in which there were two sets of 3-component correlation values, covering first and second viewings.

Using a decision tree classifier [29], we found mean classification scores across leave-one-out modeling as shown in Table 4. For the PV model, the error was 0.33 ± 0.56 (chance error = 0.66). For the NV model, the error was 0.11 ± 0.04 (chance error = 0.50). Generally, we found that whole-movie RCA-based correlation values (in fact, a sum over all components' correlation values) above a certain threshold was characteristic of Participatory film classifications, while those below were characteristic of Expository film classifications (see Figure 6 for visual representation of this, too).

As Table 4 shows when compared to Table 1, the NV classifier is able to classify nearly all movies correctly, making only one mistake ("Fight for the Planet" incorrectly classified

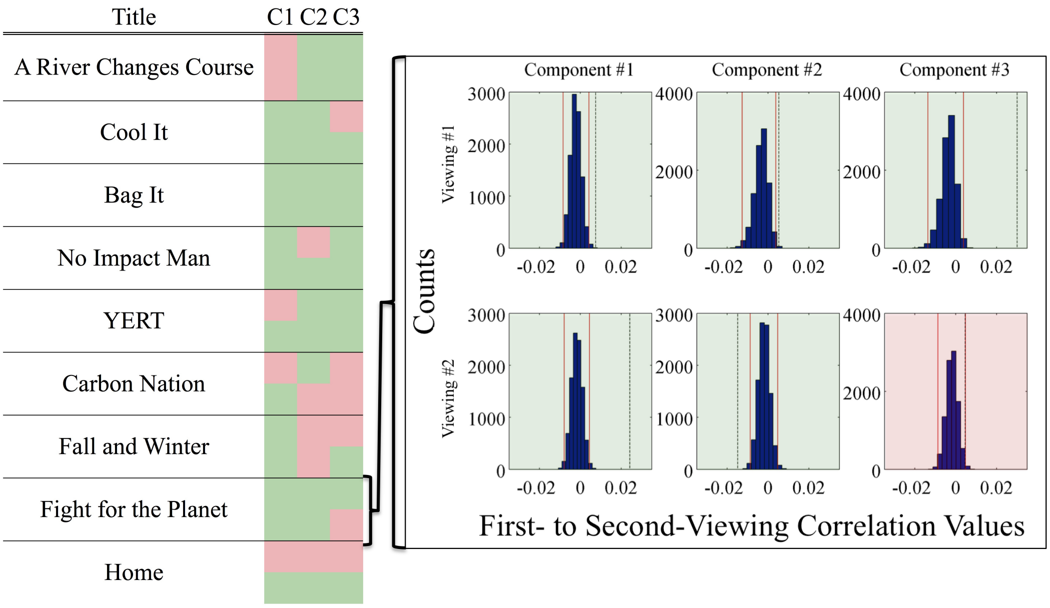


Figure 6 Overview of all movie trailers' whole-movie correlations across all subjects. Inset shows correlation values of each component and viewing against null distributions for all subjects' EEG on indicated movie trailer (e.g., Fight for the Planet). The top row indicates correlation values across subjects for the first viewing, while the bottom indicates those of the second viewing. Each column represents the indicated component number. The x-axes of each plot indicates correlation values and the y-axes indicate counts obtained from permutation testing. Blue bars indicate histograms from permutations (N = 10,000) and vertical red lines on either side of the distributions indicate p = 0.025 thresholds (two-tailed). Vertical black dashed lines indicate actual correlation values. Green and red shading of each plot indicate whether the actual correlation values are significant (green) or non-significant (red). The summary table shows significant (green) and insignificant (red) component/viewing combinations for each movie, with C1, C2 and C3 representing the first, second and third components.

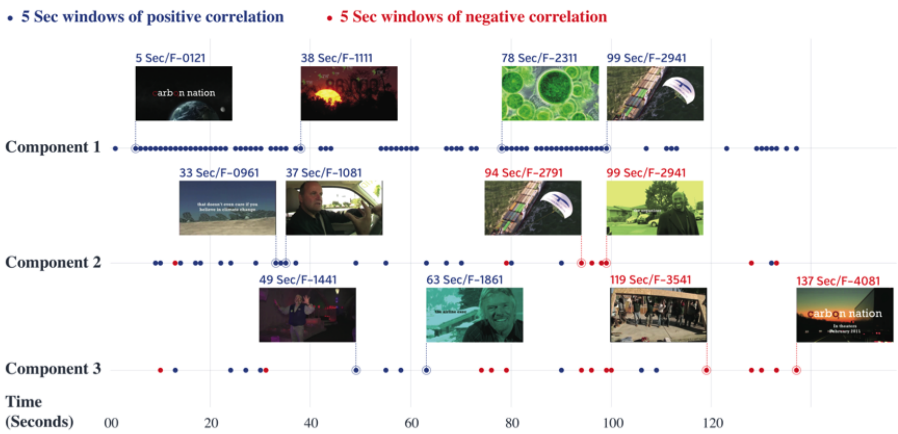


Figure 7 Example of time-resolved RCA for given subject watching a movie trailer (e.g., Carbon Nation). Each row shows 5-sec windows that are positively (blue) and negatively (red) significant from first- to second-viewing. Each indicated frame shows time in seconds (Sec) and frame number (F).

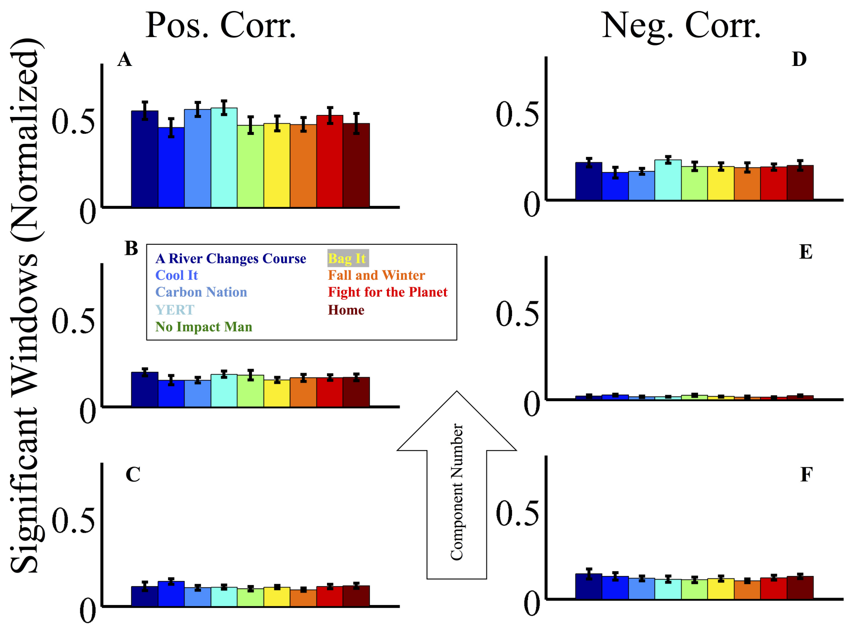


Figure 8 Number of significantly correlated 5-second windows normalized by movie length for each movie. The height of each bar indicates the mean number of windows and error bars indicate the standard error over all subjects. Each movie is color-coded according to the legend in the inset. Positive correlations of the first (A), second (B), and third (C) components are shown beside negative correlations of the first (D), second (E) and third (F) components from subjects' EEG.

more often as Participatory than Expository). When we examine this movie more closely in Table 1, we see from the algorithmic determination of NV that there was some coder disagreement as to the Narrative Voice. Although not an equal split in disagreement in the classifier score, the score for Participatory NV is not absolute (i.e., 1.00, as it is for other film trailers).

Also, comparing Table 1 and Table 4 for PV, we find significantly above chance performance for the classifier (chance accuracy = 0.33). Examining the errant classifications in Table 4, we find that the decision tree scores reflect the variability seen in the inter-coder reliability of PV, and the summarized algorithmic PV values of Table 1. For instance, as an example of a correct classification, "A River Changes Course" was classified with highest score to be a high PV (0.74), though algorithmic determination of PV is mid-high. The classification score for mid PV was the second highest for this film trailer (0.25). Alternatively, as an example of incorrect classification, "Bag It" was classified as low PV (0.90), though algorithmic determination of PV is mid-low. The classification score for mid PV was the next highest classification score (0.09). Similar trends were seen with other films.

4 Discussion

In this study, we identified neural correlates and behavioral metrics that differentiate viewer response amid a subject cohort that first watches a social issue documentary trailer and then is tasked with choosing a level of support/interest for that video's topic. This task simulates an important aspect of video media viewing, whereby consumers are able to instantly exercise behavioral response to such content via social media, video-streaming or economic means. Differences in neural and/or behavioral activity across movies that displayed a range of

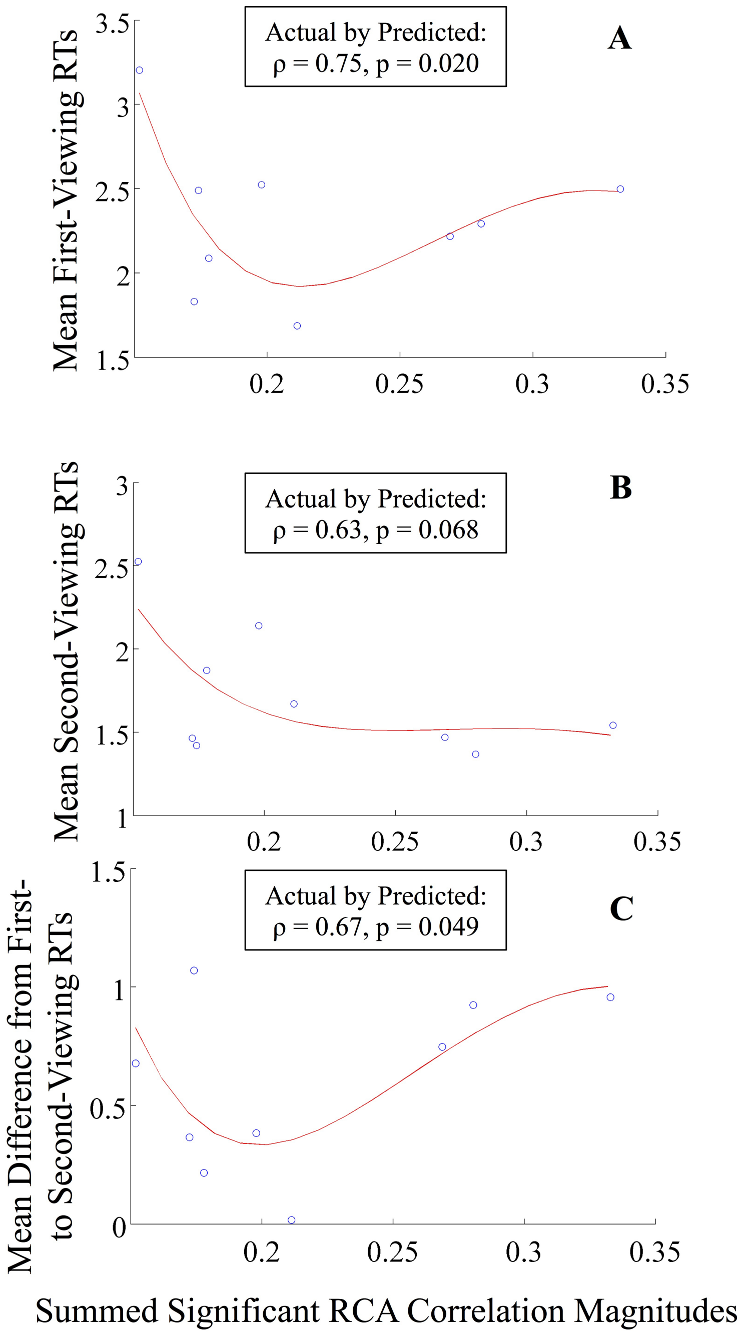


Figure 9 Variation of response times (RTs) with RCA correlation magnitudes. Blue dots represent each movie. Red lines indicate curve-fits for each plot (A-C). All curve-fits are of the form $y=Ae^{-x}+B\ln(x)+Cx+D$. Each curve-fit is shown with actual vs. predicted correlation values and p-values. Variation of first-viewing RTs (A) was found to be significantly fit (ρ = 0.75, p = 0.020). Variation of second-viewing RTs (B) showed a trend, though non-significant at our p = 0.05 threshold (ρ = 0.63, p = 0.068). Variation of the mean difference from first- to second-viewing RTs (C) was found to be significantly fit (ρ = 0.67, p = 0.049).

Table 4 Decision tree classifier outputs for Production Value (PV) and Narrative Voice (NV), using RCA whole-movie viewing metric inputs. Winning scores are highlighted in **bold**.

	Decision tree PV scores			Decision tree NV scores	
Title	Low	Mid	High	Expository	Participatory
A River Changes Course	0.01	0.25	**0.74**	**1.00**	0.00
Cool It	**0.97**	0.02	0.02	0.29	**0.71**
Bag It	**0.90**	0.09	0.01	0.15	**0.85**
No Impact Man	**0.71**	0.28	0.02	0.33	**0.67**
YERT	**0.72**	0.26	0.02	0.34	**0.66**
Carbon Nation	0.01	0.00	**0.99**	**1.00**	0.00
Fall and Winter	0.01	0.42	**0.57**	**1.00**	0.00
Fight for the Planet	**0.67**	0.31	0.02	0.02	**0.98**
Home	0.01	0.45	**0.54**	**1.00**	0.00

production values and narrative styles manifested themselves in a combination of temporally precise and whole-movie neural components, and choice value selection and response times. Furthermore, we showed that the neural components from viewing could predict above chance production value (PV) ratings and narrative voice (NV). Below, we discuss these results within the context of previous approaches to capturing neural and behavioral response to video media, especially audience reaction to movie trailers. Finally, we conclude by proposing that the system developed for this study is a viable platform to analyze production choices in movie trailer content creation.

4.1 Faster and More Extreme Behavioral Choices Relate to Social Issue Behaviors

Our behavioral observations demonstrate the tremendous variability in support choice a viewer has for a given film trailer. But we found that these choices shift towards extremes of support/non-support in the aggregate (via more donations and disinterested support choices on second-viewing) and that they are made faster upon second viewing. The increase in decision speed fits within expectations established by brand recognition literature. For instance, MacDonald and Sharp [22] replicated a classic study by Hoyer and Brown [15] in which subjects' product choices were gauged by response time for products having a brand name either previously known or unknown to the subject. They found that response times to the known brand were significantly quicker than those to the unknown, inferring that more decision effort is exercised in the case of unknown brands' product characteristics, thus causing longer decision times. Applying this model to each film trailer, the support choice decision upon second viewFigure 4ing is made faster concerning its brand (Figure 4B), due to the first viewing having occurred already, i.e., each subject has already obtained greater familiarity with the movie brand as communicated by its trailer.

A possibly unintended correlate of this faster response time is that, on a population level, the support level is pushed to the extreme. The shift in choices from Facebook 'Like' and movie-streaming 'Add to Queue' to 'Donate' and 'Not Interested' (Figure 4A) supports our hypothesis that these choices can be arranged in the cardinal order shown in Table 2. Previous work on the relationship between social media activism and economic activism supports this cardinality. For instance, Qualman coins a term socialnomics to describe the phenomenon by which social media movements are transformed into movements with economic and other consequences, rather than the reverse [26]. The relationship between 'Add to Queue' and 'Not Interested' can be understood in the context of brand recognition effects on response times. Decisions of 'Not Interested' have a lower time cost than 'Add to Queue' decisions, since the latter will require further time investment to see the movie in its entirety. Finally, the relationship between a social action (e.g., Facebook 'Like') vs. a secretive one (e.g., 'Add to Queue') can be understood in the context of secret ballot behavior in political elections. Nichter considers the influence of economic incentives in voting behavior when a ballot is cast in secret, thus ensuring a barrier between the entity purchasing the vote and the one casting it [25]. A starting point for this analysis is that voting behavior differs by whether that vote is done in private or public. For instance, in the case of a Facebook 'Like', a vote is publicly cast in support of a given movie trailer or the movie itself. Conversely, there is no public involvement in the choice to 'Add to Queue' because the decision is only recorded in a given movie-streaming account. Using the paradigm and system described in this study, the consequences of public versus private actions could be studied in further detail, though more precise manipulation of public vs. private design variables would be needed.

4.2 RCA Metrics Relate to Narrative Features Previously Thought To Be Qualitative

Our whole-movie trailer and time-resolved RCA results demonstrate that the experimental paradigm designed for this study reveals consistent neural response with a subject cohort that is relatively small in size when compared to other forms of audience testing. While we found variability in mean component correlations between first and second viewing across movie trailers, there was no significant difference in standard errors across the subject cohort (Figure 5). We found similar variability in the mean number of significant 5-second EEG correlations, but also no difference in their standard errors across the subject cohort (Figure 8). With no other participatory criteria than possession of both movie-streaming and Facebook accounts, the uniformity in within-movie neural response for such a small subject population (compared to those of market surveys and focus groups) shows that consistent neural response for a large population can be obtained from a relatively small sample size, provided that subjects are chosen within paradigm-relevant constraints. Such consistency is somewhat of a confirmation of previous results obtained by Dmochowski et al. [3], in which a subject cohort on the order of 20 subjects whose neural response was collected and analyzed were predictive of certain behavioral responses of millions of social media users.

One of the most compelling aspects of this study is the link the neural components provide to production value (PV) and narrative voice (NV) characteristics of each film trailer. For PV, we found a significantly above-chance ability to classify the quality of movie production from only the neural components across the population. Even more compelling, for NV, we found a near-perfect accuracy in classifying Expository vs. Participatory NV from whole-movie viewing RCA calculated across the population of subjects. These results indicate that a population-level neural indicator for audience perception of PV and NV may exist in the brains of viewers and, more impressively, be measurable with EEG. A possible reason for the neural basis of PV discrimination is not as simple as one for NV, but there is corroborating neural literature to provide a starting basis for explaining this result. First, for PV discrimination, viewers used in this study have had extensive exposure to the produced films that are standard fare of modern media. In that prolonged exposure, an implicit understanding of high- and low-quality production develops, as it does in music, theater and other temporal forms of media [2, 7, 32, 33, 21]. Due to this accumulated experience, the subject population implicitly is able to grade any viewed movie segment from this experiential context. The extent to which that context impacts the slow-wave variations measured in the EEG is not revealed by this paper, but it is probable that the comprehension measurable with RCA is impacted by the attention demands either mediated or inhibited by high, low or medium PV [16, 23].

The classification of NV from the whole-movie viewing RCA alone also has potential roots in current neural literature. It is widely believed that the default mode network (DMN) is linked to autobiographical narrative and self-awareness [27]. While the DMN activity is prominent in a non-viewing situation, its balance with networks involved in decision-making and active attention is also widely established [27, 20], providing an overall energy balance in activity between the various networks of the human brain. Due to the Participatory PV being a first-person telling, it is possible that the whole-movie viewing RCA incorporates a population-level measurement of DMN activity, either indirectly or directly, that is measurable with the EEG and ultimately classifiable. The finding that Participatory films tended towards having generally higher correlation values than Expository films could indicate that audience comprehension is augmented from this NV choice. While further work is needed to establish this hypothesis, for instance with MRI-based imaging and analysis methods, or expanding the number of films analyzed in this manner, the possibility of an audience-wide neural measurement of first-person engagement opens new avenues in the choice of NV available to content producers, potentially moving media/narrative creation techniques beyond current theoretical constructs and narrative classification schema (e.g., [24]).

4.3 RCA Metrics' Relation to Behavioral Decision Time May Index Movie Narrative Comprehension

Another of the most compelling aspects of this study is the potential link between behavioral and neural response. The aim of accessing nervous system response to social issue video media is to get closer measurement of the circuits executing decisions whose behavior is often unpredictable and difficult to understand. For instance, we found such variability in the utter inconsistency of choice values for a given subject and movie pairing (Figure 3). The experimental testing of our hypothesis concerning RCA-based correlation metrics and response times (hypothesized relationship: Figure 2; actual relationship: Figure 9) shows that, even with a small set of movie trailers (e.g., nine) within a given social issue (e.g., environment), a consistent relationship emerges between neural metrics previously shown to index media comprehension and behavioral metrics connected to cognitive conflict of choice. Figure 9A shows that the low RCA-based correlation, i.e., inconsistent neural response and hence possibly low video comprehension, occurs before a support choice that takes a comparatively long time to execute. Considering this result from the perspective of brand recognition and cognitive conflict [22], this relationship could mean that movie trailers for which response times are high are not communicating their message in a way that encodes into the nervous system of the viewer, leading to confusion upon presentation of the choice screen immediately following the trailer's conclusion. Furthermore, this trend remains upon second viewing (Figure 9B), hence making the average drop in RTs due to repeated viewing not strong enough to counteract the lingering choice conflict that follows viewing (see Figure 9C).

Considering previous work linking RCA-based metrics to comprehension, the relationship between medium/high RCA-based correlations and RTs could point to interplay between surety of choice and invoked understanding in the audience. First-viewing RTs rise to a plateau for high RCA-based correlations, while second-viewing RTs trend towards a sink (see Figure 9A and Figure 9B). No firm conclusions can yet be drawn on this difference without a larger stimulus set, i.e., more movie trailers, to confirm or disprove. A preliminary conclusion could be that reliable neural activity upon first viewing instigates a consideration of the appropriate level of support once the movie ends which would be bounded by the penalty of time spent on decision. The bounding creates the rise to the plateau. Such a penalty does not arise for cases of low RCA-based correlation because cognitive effort is still being spent on video comprehension. In the case of second viewing though, no such time penalty exists: comprehension is high and so anticipation of the target choice precedes presentation of the choice screen, thereby reducing RT. Finally, the mid-range RCA metrics demonstrate a possible 'sweet spot' in viewer certainty, where a sufficient balance between consistency and variability of neural response provokes the fastest RTs. Further research is needed to gauge exactly why such a valley in RTs exists as a function of RCA-based correlation metrics.

Acknowledgements. This work was supported by a grant from the Rita Allen Foundation and the Harmony Institute. Joanna Raczkiewicz contributed to experimental design discussions and Sher Chew designed the presentation of data shown in Figure 7. Julian Saliani assisted with decision tree analysis.

References

1. D. H. Brainard. The Psychophysics Toolbox. *Spatial Vision*, 1997.
2. Simon Carlile. Psychoacoustics. In *The Sonification Handbook*, pages 41–61. Logos Verlag, 2011.
3. J. P. Dmochowski, M. A. Bezdek, B. P. Abelson, J. S. Johnson, E. H. Schumacher, and L. C. Parra. Audience preferences are predicted by temporal reliability of neural processing. *Nat Commun*, 5:4567, 2014. `doi:10.1038/ncomms5567`.
4. J. P. Dmochowski, A. S. Greaves, and A. M. Norcia. Maximally reliable spatial filtering of steady state visual evoked potentials. *NeuroImage*, 2014. URL: `http://arxiv.org/pdf/1407.6110.pdf`, `doi:10.1016/j.neuroimage.2014.12.078`.
5. J. P. Dmochowski, P. Sajda, J. Dias, and L. C. Parra. Correlated components of ongoing EEG point to emotionally laden attention – a possible marker of engagement? *Front Hum Neurosci*, 6:112, 2012. `doi:10.3389/fnhum.2012.00112`.
6. M. M. Farbood, D. J. Heeger, G. Marcus, U. Hasson, and Y. Lerner. The neural processing of hierarchical structure in music and speech at different timescales. *Front Neurosci*, 9:157, 2015. `doi:10.3389/fnins.2015.00157`.
7. Hugo Fastl and Eberhard Zwicker. *Psychoacoustics: Facts and models*. Springer, 2007. `arXiv:arXiv:1011.1669v3`, `doi:10.1007/978-3-540-68888-4`.
8. J. Finsterwalder, V. G. Kupppelwieser, and M. de Villiers. The effects of film trailers on shaping consumer expectations in the entertainment industry – A qualitative analysis. *Journal of Retailing and Consumer Services*, 2012. `doi:10.1016/j.jretconser.2012.07.004`.
9. O. Furman, U. Hasson, L. Davachi, Y. Dudai, and N. Dorfman. They saw a movie: long-term memory for an extended audiovisual narrative. *Learn Memory*, 14(6):457–467, 2007.
10. U. Hasson, O. Furman, D. Clark, Y. Dudai, and L. Davachi. Enhanced intersubject correlations during movie viewing correlate with successful episodic encoding. *Neuron*, 57(3):452–462, 2008. `doi:10.1016/j.neuron.2007.12.009`.
11. U. Hasson, A. A. Ghazanfar, B. Galantucci, S. Garrod, and C. Keysers. Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends Cogn Sci*, 16(2):114–121, 2012. `doi:10.1016/j.tics.2011.12.007`.
12. U. Hasson, Y. Nir, I. Levy, G. Fuhrmann, and R. Malach. Intersubject synchronization of cortical activity during natural vision. *Science*, 303(5664):1634–1640, 2004. `doi:10.1126/science.1089506`.

13 U. Hasson, E. Yang, I. Vallines, D. J. Heeger, and N. Rubin. A hierarchy of temporal receptive windows in human cortex. *J Neurosci*, 28(10):2539–2550, 2008. doi:10.1523/JNEUROSCI.5487-07.2008.

14 Christopher J. Honey, Thomas Thesen, Tobias H. Donner, Lauren J. Silbert, Chad E. Carlson, Orrin Devinsky, Werner K. Doyle, Nava Rubin, David J. Heeger, and Uri Hasson. Slow Cortical Dynamics and the Accumulation of Information over Long Timescales. *Neuron*, 76(2):423–434, 2012. doi:10.1016/j.neuron.2012.08.011.

15 W. D. Hoyer and S. P. Brown. Effects of Brand Awareness on Choice for a Common, Repeat Purchase Product. *Journal of Consumer Research*, 17:141–148, 1990.

16 Laurent Itti. Automatic foveation for video compression using a neurobiological model of visual attention. *IEEE Transactions on Image Processing*, 13(10):1304–1318, 2004. doi:10.1109/TIP.2004.834657.

17 D. Jerrick. The Effectiveness of Film Trailers: Evidence from the College Student Market. *UW-L, Journal of Undergraduate Research*, XVI, 2013.

18 J. Kato, H. Ide, I. Kabashima, H. Kadota, K. Takano, and K. Kansaku. Neural correlates of attitude change following positive and negative advertisements. *Front Behav Neurosci*, 3:6, 2009. doi:10.3389/neuro.08.006.2009.

19 D. R. J. Laming. *Information theory of choice-reaction times.* Academic Press, 1968. doi:10.1002/bs.3830140408.

20 Baojuan Li, Xiang Wang, Shuqiao Yao, Dewen Hu, and Karl Friston. Task-dependent modulation of effective connectivity within the default mode network. *Frontiers in Psychology*, 3(JUN), 2012. doi:10.3389/fpsyg.2012.00206.

21 Weisi Lin and C. C. Jay Kuo. Perceptual visual quality metrics: A survey. *Journal of Visual Communication and Image Representation*, 22(4):297–312, 2011. doi:10.1016/j.jvcir.2011.01.005.

22 E. K. Macdonald and B. Sharp. Brand Awareness Effects on Consumer Decision Making for a Common, Repeat Purchase Product:: A Replication. *Journal of Business Research*, 2000. doi:10.1016/S0148-2963(98)00070-8.

23 M.Vranjes, S. Rimac-Drlje, and O. Nemcic. Influence of foveated vision on video quality perception. *2009 International Symposium ELMAR*, pages 28–30, 2009.

24 Bill Nichols. *Introduction to documentary.* Indiana University Press, 2001. arXiv:arXiv:1011.1669v3, doi:10.1017/CBO9781107415324.004.

25 S. Nichter. Vote Buying or Turnout Buying? Machine Politics and the Secret Ballot. *American Political Science Review*, 102(1):19–31, 2008.

26 E. Qualman. *Socialnomics: How Social Media Transforms the Way We Live and Do Business.* Wiley, 2012.

27 Marcus E. Raichle. The Brain's Default Mode Network. *Annual review of neuroscience*, 38:433–447, 2015. doi:10.1146/annurev-neuro-071013-014030.

28 Mor Regev, Christopher J. Honey, Erez Simony, and Uri Hasson. Selective and invariant neural responses to spoken and written narratives. *The Journal of Neuroscience*, 33(40):15978–88, 2013. doi:10.1523/JNEUROSCI.1580-13.2013.

29 Lior Rokach and Oded Maimom. *Data mining with decision trees: theory and applications.* Springer, 2007. doi:10.1007/978-0-387-09823-4.

30 G. J. Stephens, L. J. Silbert, and U. Hasson. Speaker-listener neural coupling underlies successful communication. *Proc Natl Acad Sci U S A*, 107(32):14425–14430, 2010. doi:10.1073/pnas.1008662107.

31 HI storypilot. https://storypilot.org/. Accessed: 2016-05-02.

32 Zhou Wang and Qiang Li. Video quality assessment using a statistical model of human visual speed perception. *Journal of the Optical Society of America. A, Optics, image science, and vision*, 24:B61–B69, 2007. doi:10.1364/JOSAA.24.000B61.

33 Stefan Winkler, Animesh Sharma, and D. McNally. Perceptual video quality and blockiness metrics for multimedia streaming applications. ... *Wireless Personal Multimedia ...*, 2001. URL: `http://stefan.winkler.net/Publications/wpmc2001.pdf`.

34 J. R. Zhang, J. Dmochowski, P. Sajda, J. R. Kender, and J. Sherwin. Correlating speaker gestures in political debates with audience engagement measured via EEG. In *ACM Multimedia 2014*, Orlando, FL, USA, 2014.