SCIENTIFIC REP🅞RTS

OPEN

Engaging narratives evoke similar neural activity and lead to similar time perception

Samantha S. Cohen[1], Simon Henin[2] & Lucas C. Parra[2]

It is said that we lose track of time - that "time flies" - when we are engrossed in a story. How does engagement with the story cause this distorted perception of time, and what are its neural correlates? People commit both time and attentional resources to an engaging stimulus. For narrative videos, attentional engagement can be represented as the level of similarity between the electroencephalographic responses of different viewers. Here we show that this measure of neural engagement predicted the duration of time that viewers were willing to commit to narrative videos. Contrary to popular wisdom, engagement did not distort the average perception of time duration. Rather, more similar brain responses resulted in a more uniform perception of time across viewers. These findings suggest that by capturing the attention of an audience, narrative videos bring both neural processing and the subjective perception of time into synchrony.

Most people are familiar with the experience of becoming completely captivated by a book or movie[1]. This phenomenon has been described as "engagement", "transportation", "absorption", or "flow"[2–4]. Some argue that when we are fully absorbed in a narrative, there is a loss of conscious awareness of the external environment and the passage of time[2, 5, 6]. Attentionally demanding stimuli typically seem to take less time[7, 8]. However, this finding may not be universally true[9]. For instance, the nature of the temporal distortion is likely dependent on the emotional valence of the stimulus[10].

Here we will explore the relationship between attentional engagement with video narratives and the perceived passage of time on the scale of seconds. Previous research on the neural basis of time perception has examined durations that range from milliseconds to days[11]. For stimuli in the order of seconds, previous work mostly concerns controlled stimuli, such as constant tones or images[12]. Given the importance of attention to time perception[8, 13–18], it is possible that time perception may be substantially different during more realistic scenarios, such as engaging narratives[19]. Attention is also known to modulate the similarity of electroencephalographic (EEG) evoked responses across viewers during narrative videos[20]. In this paper, attentional engagement was assessed from the inter-subject correlation (ISC) of EEG responses[21]. For an objective reference of stimulus engagement, we also define a behavioral measure of engagement that is based on the time that online viewers commit to watching the narrative videos. As we will show, this is an objective, value-based metric. Unlike previous measures[5, 22], it is independent of subjective self-report biases and its assessment does not interrupt the processing of the stimulus. Engagement, whether measured neurally, or behaviorally, is thus defined by the commitment to devote a scarce resource to the stimulus. In this case, that resource is either attention or time, and, as predicted, these engagement measures are correlated. Next, we address whether moments of high engagement prolong or shorten the perception of time. Surprisingly, neither the behavioral measure of time commitment, nor the neural measure of attention coincided with the perceived passage of time. Instead, the similarity of brain responses predicted the uniformity of time estimates across viewers. This robust effect was reproduced across two cohorts of viewers. Thus, engagement does not appear to distort perceived time duration, but rather, engagement, inducing a more uniform neural processing of the stimulus, leads to a more uniform assessment of time.

Results

We define engagement as the commitment to devote a scarce resource, such as attention or time, to a stimulus. Three experiments were performed to establish behavioral and neural measures of engagement, and to relate these assessments to time perception. In the first experiment, behavioral engagement was evaluated from the

[1]The Graduate Center of the City University of New York, New York, USA. [2]The City College of the City University of New York, New York, USA. Correspondence and requests for materials should be addressed to L.C.P. (email: parra@ccny.cuny.edu)
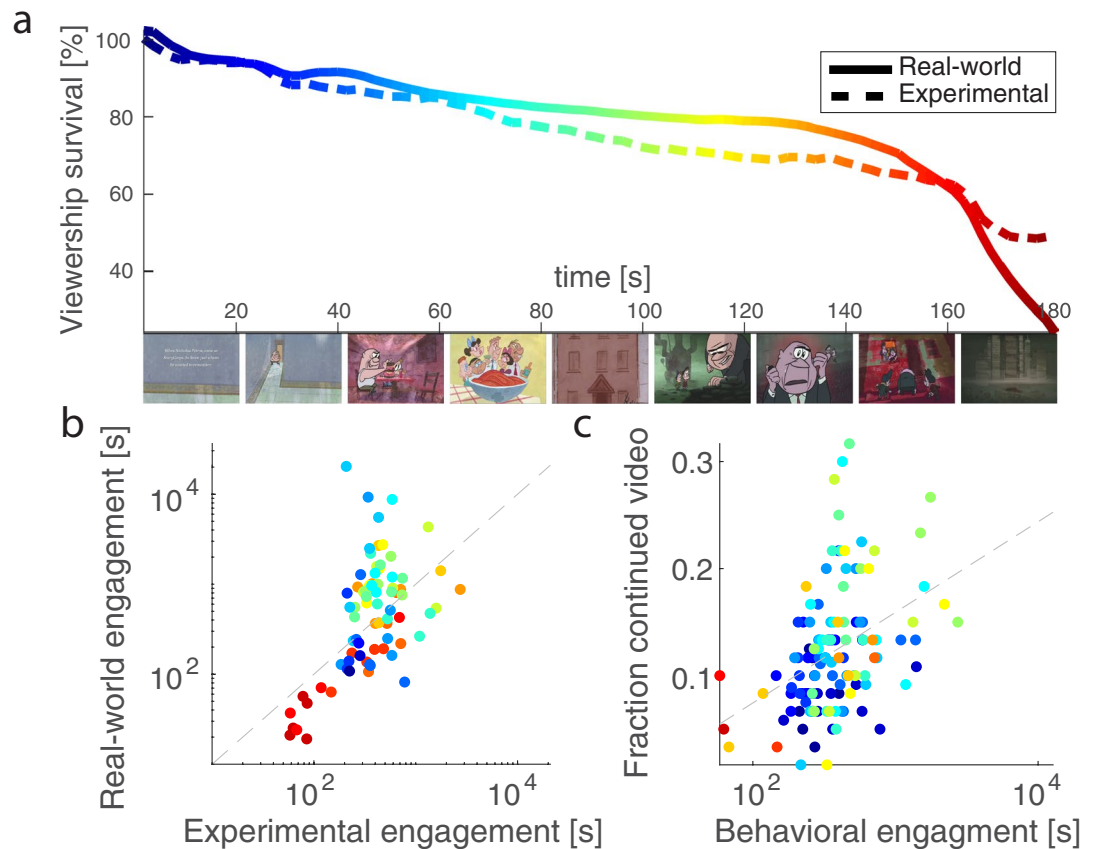
Figure 1. Behavioral engagement in "experimental" cohort mimics "real-world" behavior. (**a**) Viewership survival, shown for an example video, is measured as the fraction of the initial number of viewers who are retained over time. "Real-world" data (continuous line) was provided for this study by its content owner, StoryCorps, and represents the viewership (239,511 views) accumulated over several years for this video. "Experimental" survival (dashed line) was collected in approximately one hour from 1000 viewers recruited online via Amazon's Mechanical Turk (MTurk) platform. (**b**) Variation in engagement across time correlates between real-world and experimental data (r = 0.60, p = 0.002, N = 77). Engagement was estimated using a time interval of $\Delta t = 12$ s for the five videos that were common to both conditions. Dashed line indicates points with equal engagement in both data-sets. (**c**) Experimental engagement correlated with the likelihood that a separate cohort of viewers voluntarily continued to watch the videos when given the option to stop. Dashed line represents the regression line. In both b and c each point represents a time interval, color represents the corresponding time point in the video from a, and engagement is displayed on a log-seconds scale. Still images from "Sundays at Rocco's", a StoryCorps animated short directed by the Rauch Brothers and produced by Lizzie Jacobs and Mike Rauch, reproduced here with permission from StoryCorps. We would additionally like to acknowledge StoryCorps for providing us with their YouTube analytics data and their animated interviews as video stimuli. These interviews were recorded by StoryCorps and are provided courtesy of StoryCorps, a national not-for-profit corporation dedicated to preserve and share humanity's stories in order to build connections between people and created a more just and compassionate world. www.storycorps.org. This figure is not covered by the CC BY license. Credits to StoryCorps in panel a. All rights reserved, used with permission.

viewing behavior of large online audiences. In the second experiment, neural engagement was extracted from the EEG responses to the same videos for which behavioral engagement had been measured. In the third experiment, the perception of time was queried during short intervals within the videos.

Experimental assessment of engagement behavior.   Time commitment can be calculated from the stimulus's ability to retain viewers. For a large enough audience, this can be measured from viewership survival, $S(t)$, defined as the fraction of the audience that "survives" until time, t, in the video (Fig. 1a). The rate at which the audience shrinks is the risk of viewership loss, denoted here as $\lambda(t)$ (Equation (3) in Methods, Risk of viewership loss section, and Supplementary Fig. S1). When the stimulus evokes a high level of engagement, the risk of losing viewers is low. Conversely, when the audience is not engaged, the risk is high. Instantaneous behavioral engagement is formally defined here as the inverse of the risk of losing viewers:

$$E(t) = 1/\lambda(t). \tag{1}$$

In the present context, where time is the scarce resource being allocated, $E(t)$ is equivalent to the additional time, measured in seconds, that the average viewer is willing to invest in the stimulus (for more detail on interpretation, see Methods, Risk of viewership loss section and Discussion).

Raw viewership survival data (Fig. 1a) was collected for two cohorts of viewers. One cohort voluntarily watched the video stimuli online (real-world, 5 videos, approximately 2 million viewers over $2.6 +/- 0.8$ years, data shared by StoryCorps). The second cohort consisted of a group of subjects who were directed to the videos as part of an experimental paradigm (experimental, 10 videos, N = 1000, collected over approximately 1 hour on Amazon's Mechanical Turk platform, MTurk; see Methods, Behavioral Engagement Data Collection section). All video stimuli were animated renditions of biographical narratives (161 ± 44 s, mean $\pm$ standard deviation). Five of the videos were viewed by both cohorts. In the real-world, viewers commit time to watch videos despite real-world commitments and time pressures. In contrast, experimental viewers were given an artificial time pressure of 15 minutes to access 27 minutes of video. For the videos common to both groups, real-world viewers were found to be significantly more engaged than those recruited experimentally (t(4) = 3.2, p = 0.03, paired t-test), Supplementary Fig. S2).

Despite this disparity in overall engagement, the two groups exhibited a similar modulation in engagement over the course of the videos. This relationship was present regardless of the time scale (Δt) at which engagement was evaluated (in Equation (3)). The correlation between the experimental and real-world datasets was stable for time intervals, Δt, ranging from 4 to 21 seconds (r = 0.57 +/- 0.06). Here, and in all subsequent analyses, time and engagement were measured on a logarithmic scale following convention in time perception[23] and survival analysis literature[24]. Figure 1b displays this relationship for a time interval of $\Delta t = 12$ s. While in the real-world viewership drops by approximately 9,000 viewers in an interval of 12 s, in the experimental cohort only 7 viewers are lost. Therefore, the experimental assessment of engagement is a noisier metric (see Methods, Behavioral Engagement Data Collection section, and Supplementary Fig. S1). To validate the interpretation of this measure of engagement as a time commitment, we assessed the willingness to continue to watch the videos in a separate experiment (Results Section 3). Since this additional cohort was being compensated for performing a different task, the decision to continue to watch a video represents both a time and a financial sacrifice. Engagement measured experimentally was correlated with the fraction of viewers that voluntarily elected to continue to watch the videos after completing the time estimation task (r = 0.40, p = 0.009, N = 121, Fig. 1c). Here and in all subsequent analyses correlation p-values are computed using nonparametric phase-shuffle statistics to account for correlated data points (see Methods, Statistics section). These independent validations indicate that the experimentally derived measure of engagement is a good data-set with which to evaluate a neural measure of engagement.

Neural engagement predicts behavioral engagement.

We previously proposed that the similarity of electroencephalographic (EEG) evoked responses across viewers may be a neural marker of engagement[21]. The similarity of EEG activity across subjects can be measured as the inter-subject correlation (ISC) of stimulus evoked responses[21]. Since ISC is sensitive to attentional state[20], we predicted that there would be a relationship between ISC and the behavioral measure of engagement. To calculate ISC, EEG was recorded from 20 individuals who watched the same 10 videos from above. Components of maximal inter-subject correlation were then extracted from the EEG (see Methods, Inter-Subject Correlation section). These components, C1–C3, capture sources of the evoked neural responses that are correlated in time across the entire sample of viewers (corresponding spatial distributions shown in Fig. 2a). As such, each component potentially captures a different aspect of neural processing (e.g. visual, auditory, or supramodal processing[25]). The ISC of each component can be resolved in short time intervals during the stimulus[21, 26], and time-resolved ISC was used to predict time-varying behavioral engagement.

A regression model was first fit to the experimental behavioral engagement data (see Methods, Equation (4) Comparisons between behavioral and neural engagement section). This model's predictive ability was then tested on the real-world data. Goodness of fit was assessed for different time intervals, Δt, over which both behavioral engagement and ISC were calculated (Supplementary Fig. S3). $\Delta t = 12$ s was selected as a good compromise between performance and number of samples (i.e. this Δt had the smallest p-value, p = 2e-6 with, R = 0.4, N = 128). The predictor of engagement behavior, $\hat{E}(t)$, which we refer to as "neural engagement" can be written as a product of baseline engagement, E_0, with a time varying neural factor, $\gamma(t)$:

$$\hat{E}(t) = E_0 \gamma(t). \tag{2}$$

The baseline level of engagement, E_0, was estimated to be 212 s, and the overall estimated engagement, averaged in time for all videos, was 307 s. Thus, for the experimental behavioral data, approximately 30% of the commitment to watch the stimuli could be accounted for by the temporal variation of the neural predictor variable $\gamma(t)$ (log-sum of ISC in the largest three components; Equation (4)). Behavioral engagement was mostly explained by the ISC of the second component (C2 in Fig. 2a), which scales baseline engagement by a factor of $1.5 +/- 0.3$ (mean and std of $\gamma_2(t)$ in Equation (4), Methods, Comparisons between behavioral and neural engagement section). Components C1 and C3 of the ISC contributed relatively less ($1.0 +/- 0.005$ and $1.1 +/- 0.1$, mean and std of $\gamma_1(t)$ and $\gamma_3(t)$ respectively in Equation (4)).

To test how well this neural engagement model predicts unseen data, we compared it to the real-world behavioral engagement data and found a significant correlation (r = 0.56, p = 0.003, N = 78 intervals from 5 videos, Fig. 2b). In fact, this correlation was equally strong when training the regression coefficient with the experimental data from the five videos that were not part of the real-world behavioral data (r = 0.58, p = 0.003, N = 78). Thus, the predictive neural engagement model not only generalizes to unseen data, but it also generalizes across different stimuli.
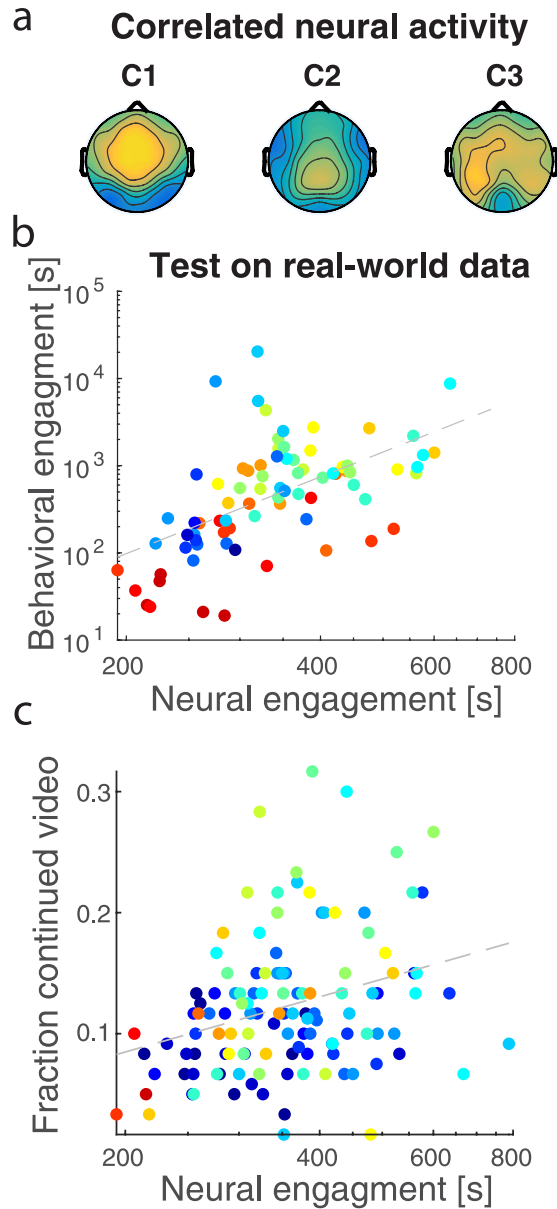
Figure 2. Neural Engagement predicts Behavioral Engagement. (**a**) Spatial distribution of the three EEG components with maximal inter-subject correlation (ISC; C1–C3). Color indicates positive (yellow) or negative (blue) correlation between the source of the neural responses and each sensor on the scalp. Component C2, center, contributes most to the relationship between neural and behavioral engagement (see main text). (**b**) Neural engagement ($\hat{E}(t)$; dashed line) predicts real-world behavioral engagement, $E(t)$ using the model developed on the experimental behavioral engagement data ($r = 0.56$, $p = 0.003$, $N = 78$, for the 5 testing videos, $\Delta t = 12$ s). (**c**) Neural engagement correlated with the fraction of a separate cohort of viewers that decided to continue to watch the videos when given the option to stop prematurely. Dashed line represents the regression line ($r = 0.31$, $p = 0.0006$, $N = 122$ intervals from all 10 videos). Each point represents a different time interval in each video colored according to time as in Fig. 1. Both engagement measures are displayed on a log-seconds scale.

Engagement behavior is relatively consistent across the time course of all videos (Supplementary Fig. S1). To test whether the relationship between neural and behavioral engagement is due to this general trend, the analysis was repeated with the engagement measures taken from different clips (the identity of the clips were shuffled prior to analysis). The resulting correlation values were significantly lower ($r = 0.28 +/- 0.1$ compared to $r = 0.56$, from 10^6 random video pairings across all 10 videos, Methods, Comparisons between behavioral and neural engagement section, $p = 7e-5$). This suggests that the neural measure of engagement captures time-varying engagement that depends on the content of the videos, rather than a common trend across videos.

As with the behavioral engagement measures, the neural engagement measure, assessed across all videos, also correlated with the voluntary election to continue to watch the videos during the time estimation task ($r = 0.31$,
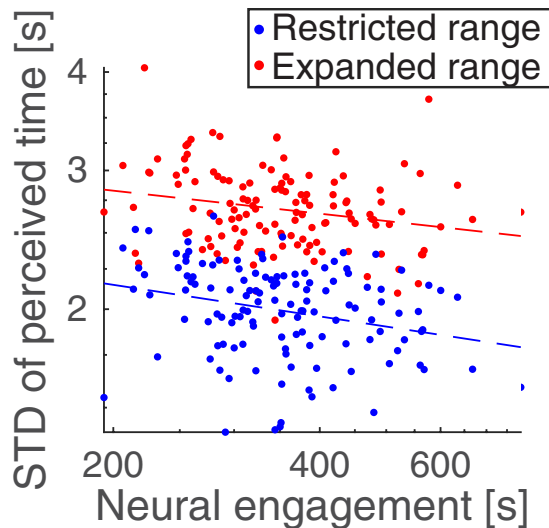
Figure 3. Engagement predicts the variability of time perception. Viewers estimated the duration of time intervals within each video. Each point represents a time interval in a video. (**a**) The duration of intervals with high neural engagement were perceived less variably across viewers ($r = -0.27$, $p = 0.0009$, and $r = -0.23$, $p = 0.05$, N = 129 intervals for all 10 videos, for restricted and expanded range, respectively). Comparisons are made across two independent cohorts who had either a restricted range (blue, 8–16 s) or expanded range (red, 4–20 s) available for their time duration estimates. All time measures are displayed on a perceptual log-seconds scale.

$p = 0.0006$, N = 122 intervals from all 10 videos, Fig. 2c). The consistency between the neural and behavioral measures confirms the hypothesis that the similarity of brain responses captures attentional engagement.

Relationship between engagement and time perception. After establishing the validity of both the behavioral and neural measures of engagement, the relationship between stimulus engagement and time perception was assessed. An additional cohort of viewers (recruited from MTurk) provided subjective estimates for the durations of brief periods of time during the videos. These segments corresponded to those for which behavioral and neural engagement were assessed ($\Delta t = 12$ s, Figs 1 and 2). Each video was shown until an interval of interest, indicated by a visual cue, and after the time interval had elapsed, subjects were asked to report the perceived duration within a range of 8 s to 16 s ("Restricted range", see Methods, Perceived Time Data Collection section for implementation details).

Consistent with existing literature[9], the duration of the 12 s intervals was underestimated ($11.3 +/- 0.03$ s). Consistent with the idea that expected events are perceived as longer due to an increase in the attention devoted to time[27–29], time intervals later in the videos were perceived as lasting longer than earlier intervals ($r = 0.57$, $p = 1e-6$, Supplementary Fig. S4). This is in conflict with the idea that context shifts in the narrative may have disrupted this effect[30–32].

Despite prevalent theories that stimulus engagement induces time distortion[22], there was no correlation between the mean estimates of time duration and engagement, measured either behaviorally or neurally ($p > 0.3$). Interestingly, however, neural engagement correlates with the *variability* of time estimates across viewers (Fig. 3, Restricted range, $r = -0.27$, $p = 0.0009$, N = 129). The same is not universally true for behavioral engagement, where the effect was weaker (for experimental behavioral engagement: $r = -0.20$, $p = 0.03$, N = 128 intervals for all 10 videos, and for real-world behavioral engagement: $r = -0.25$, $p = 0.2$, N = 78 intervals from the 5 real-world videos). Variability is measured as the standard deviation of the time estimates across viewers. Thus, people are not losing track of time, per se, but rather they are tracking time more similarly when they are engrossed in the story.

To demonstrate the reproducibility of this finding, and because the distribution of time estimates was truncated in the initial experiment (see Supplementary Fig. S5), a second cohort was recruited. This cohort reported duration within an expanded range of 4 s to 20 s ("Expanded range"). The variability of time estimates across viewers again correlated with neural engagement (Fig. 3, Expanded range, $r = -0.23$, $p = 0.05$, N = 129). Unlike the relationship between neural and behavioral engagement, it was not determined whether the relationship between time estimates and engagement was a general property of the kind of narrative videos tested, or due to the specific content of each story. See Methods, Relationship between time perception and engagement section, for additional comparisons and controls. This independent cohort largely confirms our main and novel finding that engagement synchronizes neural activity across brains, thus resulting in a more uniform perception of time across people.

Discussion

It is often said that we lose track of time when absorbed in an engaging narrative[5, 6]. Indeed, engagement can potentially either shorten or elongate our subjective perception of time[7, 9, 33, 34]. Engagement with a narrative may do both depending on its emotional valence[10]. The precise neural processing that results in time's distortion

during naturalistic narrative stimuli is a matter of ongoing exploration[19, 35]. To determine the dependence of time perception on engagement, we first characterized engagement behaviorally in terms of how an audience is retained by narrative videos (Fig. 1). To assess attentional engagement, we measured the inter-subject correlation (ISC) of stimulus-evoked EEG responses to the same videos (following[20, 21]). We found that ISC is predictive of the time that viewers were willing to spend with the videos (Fig. 2). Surprisingly, neither behavioral nor attentional engagement correlated with the perceived duration of intervals of time during the videos. Instead, when the videos were more engaging, they were processed in the brain more uniformly (resulting in higher ISC), and the perception of time was more uniform across viewers (Fig. 3). Thus, rather than losing track of time, viewers have a more consistent perception of time when highly engaged in the stimulus, at least for narrative videos on the scale of seconds.

Despite extensive research on the neural underpinnings of time perception, no consensus exists on how humans estimate time intervals[36]. There is even disagreement among the models of time perception as to whether there is an internal timekeeping mechanism[36, 37], as posited by the pacemaker-accumulator model[38] and attentional-gate model[39], or whether time judgments are made by comparing the memories of past events[40, 41]. In either case, while it is likely that different mechanisms support the perception of different time scales[11], there is also evidence for shared neural processes that implicate a diverse set of brain regions depending on task specifics[23]. Previous fMRI studies have explored the dependence of time estimation signals on attention[42], emotional content[43], and salience[44]. It is likely that all of these features altered time perception during our video stimuli.

Here we investigated the relationship between prospective time estimates on the order of seconds and engagement with naturalistic videos. Prospective estimates are those made when the subject is aware that they will be required to report on time and are generally more accurate[9, 13, 32]. In contrast, retrospective estimates are used when people are unaware of an upcoming time query, and are therefore more common in daily life[32]. Here, prospective estimates were used after subjects practiced estimating time and received feedback to increase the accuracy of these reports. The time estimation task usually began after much of the video had been viewed, indeed, subjects were similarly engaged during this divided attention task as they were during more naturalistic viewing conditions. The engagement measured from the first two experiments correlated with the decision to continue to watch the videos (in the third experiment, Figs 1c and 2c). Given that subjects were engaged in the videos, it is therefore surprising, that there was no relationship between engagement and the average time estimates. Engaging with an alternate task typically decreases prospective time estimates because the task distracts attention away from time[8, 14–16, 45–47]. It is possible that naturalistic stimuli, like those used here, might distract attention away from time in a different way than tasks directed by the experimenter. For instance when music is presented, Waldum and Sahakyan (2013) found that intervals were estimated as longer when more songs were remembered. The use of naturalistic stimuli in prospective time estimates is under-explored and may have different effects on perceived time than those found previously.

We found that more similar neural processing across viewers led to a more uniform report of time's passage. This suggests that stimulus processing may ultimately provide (or at least modulate) the input to the time integration circuit. This integration circuit may either be one driven by an internal clock, such as that posited in the attentional-gate model[39], or one based on comparing the memories of past events[40]. Either way, more uniform stimulus processing likely aligns both attentional and memory resources, leading to more uniform estimates of time across viewers. This is in contrast to models of time estimation wherein estimates are driven by an internal, stimulus-independent "clock"[48]. Similarly to findings in previous studies using naturalistic stimuli[19, 35], it may be that processing of the stimulus determines the way that time is perceived.

It is well established that time perception is affected by attention[8, 13–18, 47]. We therefore leveraged EEG responses evoked by the stimuli to gauge attentional engagement using ISC. The brain activity most predictive of engagement was the second component of the EEG activity correlated across viewers (see Fig. 2a for scalp topography). With a factor of $\gamma_2 = 1.5$, this component explains the difference in time commitment from the baseline level of approximately 200 s to the average engagement level of approximately 300 s. Interestingly, among the three strongest components of the ISC, this component is also the best at discriminating between attentional states[20]. This component is also uniquely evoked during narratives with sound, as opposed to those with only visuals, and it may therefore be representative of auditory processing[25], or potentially how strongly the stimulus' auditory narrative captures attentional resources.

The concept of "engagement" is used in a variety of contexts: a client engages a law firm, a couple is engaged to be married, or a student is engaged in the classroom. In all of these scenarios there is a commitment of either time or financial resources. Similarly, in this paper, we define engagement as the commitment of a scarce resource. Unlike self-report assessments[5, 22], this definition is quantifiable in strict numerical terms and can be applied even when subjective reports are impossible[49]. Whether the resources engaged are time or attention, a value-based decision is consistently assessed in which the value gained from consuming the narrative is compared to that of possible alternatives[50].

Our definition of engagement, in terms of devoting a scarce resource, could be translated to other media that engage people such as books, music, virtual reality, gaming, or to stimulating activities such as painting or playing sports. Time does not necessarily have to be the resource that is sacrificed. By making a commitment, viewers may be foregoing monetary compensation or social rewards to gain access to entertainment. The worth of engagement may be computed using a currency that can be traded for these and other scarce resources. We predict that if viewers are similarly entrained by the stimulus (or activity), thus eliciting a high level of ISC, they will be immune to extrinsic costs such as the time or money that they are sacrificing for the current moment's enjoyment. Their perception of time, one of the many valuable resources that they are sacrificing, will thus be driven by the stimulus, and consistently so across viewers.

Methods

Stimuli. Stimuli were chosen on the basis of their highly emotive content and the availability of online viewership data. We used these same stimuli in a previous EEG study on incidental memory[25]. They consisted of 10 different videos (5 from the New York Times' Modern Love episodes: "Broken Heart Doctor", "Don't Let it Snow", "Falling in Love at 71", "Lost and Found", and "The Matchmaker", and 5 from StoryCorps' animated shorts: "Eyes on the Stars", "John and Joe", "Marking the Distance", "Sundays at Rocco's" (depicted in Fig. 1a), and "To R.P. Salazar with Love"). Stimuli were 161 ± 44 s in duration (mean and standard deviation across stimuli) with a total duration of 26 min 48 s.

Behavioral Engagement Data Collection. Real-world viewing behavior was measured using the pool of viewers who found the five StoryCorps videos organically on YouTube, via the StoryCorps website (storycorps. org/animation), or another linked website. Anonymous YouTube Analytics data were provided as aggregated viewership survival data (Fig. 1a, Supplementary Fig. 1) by StoryCorps with permission for analysis and publication. The viewership data captured the behavior of viewers, totaling 2,528,897 across all five videos, amassed since the videos were made available online until the time of data retrieval (2.6 +/− 0.9 years).

Behavioral engagement was measured experimentally on an independent set of 1,000 subjects collected on Amazon's Mechanical Turk (MTurk) platform (requester.mturk.com). Participants with IP addresses located in the USA had access to all 10 stimuli, presented as randomly ordered embedded videos. They had access to the videos 15 minutes and were informed that the total duration of all videos exceeded the time allotted. Subjects were free to select a video by clicking on it and could watch each video for a duration determined at their discretion. Watching the videos was optional (required a mouse click) and it is therefore possible that a fraction of these participants did not watch any of the videos. Subjects were compensated after 15 minutes had elapsed regardless of whether or not they watched the videos. It took approximately one hour for this data to be collected for all 1000 subjects. This collection method was selected after pilot testing in three smaller cohorts of 100 subjects each. In the first pilot, participants had the option of immediate remuneration and therefore spent little time with the videos. In the second pilot, participants were not informed that they would be paid. In the third pilot, subjects were informed of their impending payment. The results of this pilot agreed best with the real-world data, and this format was therefore used for the final data collection. None of the pilot data was included in the final analysis reported here.

The YouTube Analytics API was used to extract the fraction of viewers (number of current viewers/total number of initial viewers) who watched each video interval. The resulting curve can be considered the viewership survival, $S(t)$, although it is not a strict survival metric because YouTube allows users to rewind and skip sections and therefore $S(t)$ sometimes increases. Note that the YouTube Analytics API divides each video into 100 sampling intervals, regardless of video duration, and was resampled to correspond with the absolute time elapsed, as described below.

Risk of viewership loss. At each time interval, a decision is made regarding whether the current activity provides more reward than another activity. In aggregate, over the population of viewers, this is reflected in the survival function, $S(t)$ the ratio of people that are still watching, or have survived, until time t. A typical example of this from the real-world data is shown in Fig. 1a. The risk, or hazard, of losing viewers, $\lambda(t)$, can be estimated from $S(t)$ and is conventionally defined as the relative change of the survival in a time interval Δt[51]:

$$\lambda(t) = -\frac{1}{S(t)}\frac{\delta S(t)}{\delta(t)} \approx \frac{1}{S(t)}\frac{S(t) - S(t + \Delta t)}{\Delta t}. \tag{3}$$

In a realistic scenario, stimuli are selected, engaged with, and finally abandoned when audience members determines that their limited temporal resources are better spent elsewhere. Figure 2a shows several real-world examples of how the number of surviving viewers, $S(t)$ decays over time. The hazard curve, $\lambda(t)$ derived from the survival function (Supplementary Fig. S1) presents a typical "bathtub" curve with high risk of viewership loss at the beginning and ending and low risk during the intervening time. In failure analysis inverse risk represents the mean time to failure[24], and for constant risk, the survival curve is exponentially decaying with the inverse risk as its time constant.

Risk of losing the audience, $\lambda(t)$, was computed according to Equation (3) on the 100 samples in the raw data. For visualization purposes, Fig. 1a and Supplementary Fig. S1 show survival, $S(t)$, and risk, $\lambda(t)$, as a function of absolute time (scaled to seconds). For the purpose of comparing experimental and real-world data, $\lambda(t)$ was resampled to a time scale of $\Delta t = 12$ s. Engagement, $E(t)$, derived from $\lambda(t)$ in Equation (1), is plotted at this scale (Supplementary Fig. S1). This time scale was motivated by the comparison with the neural data. To calculate the average number of viewers lost in a time interval of $\Delta t = 12$ s, the derivative of $S(t)$ was scaled by the initial number of viewers and then multiplied by 12 seconds.

Time intervals with negative engagement (due to rewinding and skipping on YouTube) are ignored because the logarithm of engagement is used in all analyses. Therefore, at a sampling interval of $\Delta t = 12$ s, 1 interval is excluded from the experimentally acquired data, and 3 intervals are excluded from the real-world data. This exclusion effects the comparison between the two behavioral engagement measures (Fig. 1b, a reduction from N = 81 intervals to N = 77), the comparison with the decision to continue watching (Fig. 1c, a reduction from N = 122 to N = 121), the comparison with neural engagement (Fig. 2b, Results Section 2, for the experimental cohort a reduction from N = 129 to N = 128, and for the real-world cohort a reduction from N = 81 to N = 78), and the comparison with the time estimates (reductions the same as those for neural engagement).

Perceived Time Data Collection.   Perceived time was measured on two additional MTurk cohorts. Each participant watched all 10 stimuli and a pseudo-randomly selected 12 second time interval was denoted by the appearance of a red hourglass in the corner of the video. After the interval had expired, the video was paused, and participants were asked to estimate the duration for which the hourglass had appeared (perceived time). The longest video clips had 19 intervals. To ensure that each of these intervals had at least 20 estimates, 380 subjects were recruited. Across all videos, there were N = 129 time intervals, of these 129, N = 81 corresponded to intervals in the videos for which real-world behavioral engagement was assessed. In this first cohort, subjects reported perceived duration on a visual analog scale with values ranging from 8 to 16 seconds. A histogram of reported time estimates shows a large number of responses at the boundaries of this range (see Supplementary Fig. S5, restricted range), suggesting that inputs may have been restricted. A second cohort was therefore recruited and given the option to input estimates ranging from 4 to 20 seconds (expanded range). To ensure a robust estimate of the standard deviation of perceived time across subjects, this cohort had 720 participants. After each interval had transpired, participants were also asked a comprehension question related to the content of the story in that interval (same questions as in ref. 25; mean accuracy level 84.0% +/− 37.7% across questions). After the response was recorded, participants were given the option to finish the video. When the interval that was just watched was the last interval in the video, the option to finish the video was not given. As this occurred in 7 cases, this resulted in a reduction from N = 129 intervals to N = 122 intervals. Prior to the presentation of the 10 experimental stimuli, participants were briefly acquainted with the task on four non-experimental videos. For the first video, subjects are told ahead of time that the interval is 12 seconds in duration; for the next three, interval durations are selected at random to be either 8, 12 or 16 seconds. Without prior knowledge of the duration, subjects are asked to estimate it and are subsequently informed of the correct duration. Informed consent was obtained from all participants, experiments were performed in accordance with relevant guidelines and regulations, and all data collection procedures were approved by the Institutional Review Board of the City University of New York.

Electroencephalographic Data Collection and Preprocessing.   Electroencephalographic (EEG) data were collected for a previous study from a cohort of 20 individuals who watched the same audiovisual stimuli (AV condition, in ref. 25). For more details regarding participants, EEG data collection, and preprocessing see ref. 25.

Inter-Subject Correlation.   Inter-subject correlation (ISC) is calculated by first finding linear combinations of electrodes that are maximally correlated between subjects[21]. We refer to these as correlated components, akin to principal or independent components. By construction, these linear combinations of electrodes are common to all subjects. The EEG data for each subject is then projected into this component space. Inter-subject correlation is measured in each component, by averaging the Pearson correlation coefficient of the projected time-courses between all pairs of subjects. The ISC calculation implemented here is the same as what has been published previously[20, 25]. Following previous research, the top three components of the EEG are extracted that capture maximally correlated responses between subjects[21]. To determine the correspondence between behavioral engagement and ISC, ISC was resolved in time, using the same time intervals over which behavioral engagement was measured (see next section). By resolving in time we mean that ISC between time-courses was calculated in a window of time of duration Δt. This time resolved ISC in the i-th correlated component is here denoted as $x_i(t)$.

Comparisons between behavioral and neural engagement.   A proportional hazard model[52] was used to relate engagement behavior and ISC. A logarithmic regression model is necessary because engagement, derived from viewership survival, fits best into a canonical log-risk model from the survival literature[53] and regression is an important step to determine the contribution of each component of ISC component ($x_i(t)$) to behavioral engagement. This resulted in a regression of engagement, $E(t)$, with a time dependent covariate, $\gamma(t)$, and a constant baseline engagement, E_0 (Equation (2) in Results). Following the traditional form of the proportional hazard model[24, 52], the time dependent covariate, $\log \gamma(t)$ equals the weighted sum of the predictor variables:

$$\log \gamma(t) = \sum_{i=1}^{3} \beta_i x_i(t) = \sum_{i=1}^{3} \log \gamma_i(t). \tag{4}$$

The contribution of ISC in each component, $x_i(t)$, to the total engagement can thereby be assessed from its corresponding $\gamma_i(t)$ value. Optimal model parameters β_i and E_0 were found as the best linear fit for the log-engagement data collected from the experimental group. This fit was performed separately for different time intervals (Δt). Engagement, $E(t)$, was resampled from the 100 samples in the raw data into various time resolutions (Δt = 1 s, 2 s, 3 s,… and up to 30 s) and time-resolved ISC $x_i(t)$ was also calculated on the matching time intervals. Goodness of fit, R, is shown in Supplementary Fig. S3 as a function of Δt. For consistency, all subsequent analyses were performed at 12 s resolution, which provided the best fit (see Results).

Final parameter estimates were assessed by training on the experimentally measured engagement data and testing on the real-world data without further parameter adaptation. This captures the generalization performance from trained to unseen data, as well as generalization from the experimental procedure to the real-world behavioral engagement data. To capture the generalization between different stimuli, in an additional analysis, training of the model parameters, β_i, and E_0, was limited to only those videos that were not in the testing set, i.e. training was performed on the experimental data from the 5 New York Times videos and the real-world data from the 5 StoryCorps videos was used for testing.

Finally, to assess whether this relationship was video-dependent or due to the canonical timecourse of the hazard function from which behavioral engagement is derived (see Supplementary Fig. S1), behavioral and neural engagement measures were paired between random videos and correlation values were calculated 10⁶ times from these random pairings. For all correlations, the behavioral engagement measured from one video was resampled to the match the length of the neural engagement measured from another video. Parameters were then found by

training with the experimental data, and correlation values were calculated from the relationship found using these parameters on the real-world data. The significance level was calculated by finding the proportion of shuffled data that yielded a larger magnitude correlation value than the one found in the original data.

Relationship between time perception and engagement. Correlation was assessed between both engagement measures (behavioral engagement: $\log(E(t))$ and neural engagement: using the optimal fit for β_i at $\Delta t = 12\,\text{s}$) and the perceived time of each interval, collected from two independent cohorts of subjects. Mean time perception was longer for later intervals in the videos (Supplementary Fig. S4). This drift in mean time perception with wait time does not affect the standard deviation of time estimates because the standard deviation, by definition, does not depend on the mean. Linear regressions were used to relate the standard deviation of time estimates to both neural and behavioral measures of engagement. Although not significant, for the time estimates from the expanded range cohort, the relationship between time estimate variability and behavioral engagement follows the trend found in the other cohort (r = −0.09, p = 0.4, N = 128 intervals from experimental behavioral engagement, and r = −0.25, p = 0.2, N = 78 intervals from real-world behavioral engagement). Additionally, all of the reported results are also obtained when the variability of time estimates across subjects is accounted for. To control for subject-specific biases, we subtracted each subject's mean time estimate across trials from all of their guesses. This analysis reproduces the results in Fig. 3: for the Restricted range, r = −0.29, p = 0.0008, N = 129, and for the Expanded range: r = −0.23, p = 0.008, N = 129.

Statistics. The significance of reported correlation values, r, were computed using nonparametric phase-shuffle statistics. These Nonparametric statistics are reported to rule out that temporal correlation between samples bias these results, in other words, to account for potential co-dependencies between data-points. Specifically, significance levels were computed using phase shuffled data (following[54]), which preserve the correlation structure in time but alters the time course of a temporal sequence. Therefore, N = 10^6 phase shuffled surrogates were produced. In Figs 1 and in 2b experimental behavioral engagement was shuffled. In Figs 2c and in 3 neural engagement was shuffled. Correlation coefficients were computed between each surrogate and the other variable in the analysis. Bootstrap p-values are calculated as the fraction of shuffles with correlation values more extreme than those obtained with the original time sequences. All bootstrap p-values are comparable to the values found using parametric statistics.

In the analysis of the viewer's decision to continue to watch the videos ("Fraction continued video" in Figs 1c and 2c), data from the two time estimation cohorts (Restricted and Expanded) was combined since the decision to continue watching did not vary across the cohorts. All analysis of time and engagement were performed on a log-seconds scale (except for the computation of the mean and standard deviation of the time estimates, Fig. 3). All statistical tests were performed in MATLAB (MathWorks, Natick, MA, USA).

References

1. Barthes, R. & Duisit, L. An Introduction to the Structural Analysis of Narrative. *New Lit. Hist.* **6**, 237–272 (1975).
2. Green, M. C. & Brock, T. C. The role of transportation in the persuasiveness of public narratives. *J. Pers. Soc. Psychol.* **79**, 701–721 (2000).
3. Tellegen, A. & Atkinson, G. Openness to absorbing and self-altering experiences ('absorption'), a trait related to hypnotic susceptibility. *J. Abnorm. Psychol.* **83**, 268–277 (1974).
4. Csikszentmihalyi, M. *Flow: The psychology of happiness*. (Harper & Row, 1992).
5. Busselle, R. & Bilandzic, H. Measuring Narrative Engagement. *Media Psychol.* **12**, 321–347 (2009).
6. Green, M. C. Transportation into narrative worlds: The role of prior knowledge and perceived realism. *Discourse Process.* **38**, 247–266 (2004).
7. Zakay, D. In *Time and human cognition* (eds Levin, I. & Zakay, D.) (Elsevier, 1989).
8. Macar, F., Grondin, S. & Casini, L. Controlled attention sharing influences time estimation. *Mem. Cognit.* **22**, 673–686 (1994).
9. Block, R. A. & Zakay, D. Prospective and retrospective duration judgments: A meta-analytic review. *Psychon. Bull. Rev.* **4**, 184–197 (1997).
10. Angrilli, A., Cherubini, P., Pavese, A. & Manfredini, S. The influence of affective factors on time perception. *Percept. Psychophys.* **59**, 972–982 (1997).
11. Buhusi, C. V. & Meck, W. H. What makes us tick? Functional and neural mechanisms of interval timing. *Nat. Rev. Neurosci.* **6**, 755–65 (2005).
12. Penney, T. B. & Vaitilingam, L. in *Psychology of time* (ed. Grondin, S.) 261–294 (Emerald Group, 2008).
13. Grondin, S. Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Atten. Percept. Psychophys.* **72**, 561–582 (2010).
14. Hicks, R. E., Miller, G. W. & Kinsbourne, M. Prospective and retrospective judgments of time as a function of amount of information. *Am. J. Psychol.* **89**, 719–730 (1976).
15. Block, R. A. In *Time, action and cognition* 141–152 (Springer Netherlands, 1992). doi:10.1007/978-94-017-3536-0_16.
16. Zakay, D. Attention allocation policy influences prospective timing. *Psychon. Bull. Rev* **5**, 114–118 (1998).
17. Block, R. A., Hancock, P. A. & Zakay, D. How cognitive load affects duration judgments: A meta-analytic review. *Acta Psychol. (Amst).* **134**, 330–343 (2010).
18. Brown, S. W. In *Psychology of time* 111–138 (2008).
19. Lositsky, O. *et al.* Neural pattern change during encoding of a narrative predicts retrospective duration estimates. *Elife* **5** (2016).
20. Ki, J., Kelly, S. & Parra, L. C. Attention strongly modulates reliability of neural responses to naturalistic narrative stimuli. *J. Neurosci.* (2016).
21. Dmochowski, J. P., Sajda, P., Dias, J. & Parra, L. C. Correlated components of ongoing EEG point to emotionally laden attention – a possible marker of engagement? *Front. Hum. Neurosci.* **6**, 1–9 (2012).
22. Nakamura, J. & Csikszentmihalyi, M. The Concept of Flow Optimal Experience and Its Role in Development. 89–105 (2002).
23. Merchant, H., Harrington, D. L. & Meck, W. H. Neural Basis of the Perception and Estimation of Time. *Annu. Rev. Neurosci.* **36**, 313–36 (2013).
24. Kalbfleisch, J. D. & Prentice, R. L. *The statistical analysis of failure time data*. (John Wiley & Sons, Inc., 2002).
25. Cohen, S. S. & Parra, L. C. Memorable audiovisual narratives synchronize sensory and supramodal neural responses. *eNeuro*, doi:10.1523/ENEURO.0203-16.2016 (2016).
26. Dmochowski, J. P. *et al.* Audience preferences are predicted by temporal reliability of neural processing. *Nat. Commun.* **5**, 1–9 (2014).

27. Cahoon, D. & Edmonds, E. M. The watched pot still won't boil: Expectancy as a variable in estimating the passage of time. *Bull. Psychon. Soc.* **16**, 115–116 (1980).
28. Grondin, S. & Rammsayer, T. Variable foreperiods and temporal discrimination. *Q. J. Exp. Psychol. A.* **56**, 731–765 (2003).
29. Jones, M. R. & Boltz, M. Dynamic Attending and Responses to Time. *Psychol. Rev.* **96**, 459–491 (1989).
30. Block, R. A. Temporal judgments and contextual change. *J. Exp. Psychol. Learn. Mem. Cogn.* **8**, 530–544 (1982).
31. Block, R. A. In *Time, mind, and behavior* 169–178 (Springer Berlin Heidelberg, 1985).
32. Brown, S. W. & Stubbs, A. D. The psychophysics of retrospective and prospective timing. *Perception* **17**, 297–310 (1988).
33. Avni-Badad, D. & Ritov, I. Routine and the perception of time. *J. Exp. Psychol. Gen.* **132**, 543–50 (2003).
34. Ornstein, R. E. *On the experience of time*. (Penguin, 1969).
35. Waldum, E. R. & Sahakyan, L. A role for memory in prospective timing informs timing in prospective memory. *J. Exp. Psychol. Gen.* **142**, 809–26 (2013).
36. Wittmann, M. Representation of Duration. *Nat. Rev. Neurosci.* **14**, 217–223 (2013).
37. Matell, M. S. & Meck, W. H. Neuropsychological mechanisms of interval timing behavior. *BioEssays* **22**, 94–103 (2000).
38. Treisman, M. Temporal discrimination and the indifference interval. Implications for a model of the 'internal clock'. *Psychol. Monogr.* **77**, 1–31 (1963).
39. Zakay, D. & Block, R. A. An Attentional Gate Model of Prospective Time Estimation. *Time Dyn. Control Behav.* 167–178 (1995).
40. Block, R. A. In Time and mind II: Information processing perspectives 41–59 (2003).
41. Block, R. A. & Grondin, S. Timing and time perception: A selective review and commentary on recent reviews. *Front. Psychol.* **5**, 1–3 (2014).
42. Coull, J. T., Vidal, F., Nazarian, B. & Macar, F. Functional Anatomy of the Attentional Modulation of Time Estimation. *Science* **303**, 1506–1508 (2004).
43. Dirnberger, G. *et al.* Give it time: Neural evidence for distorted time perception and enhanced memory encoding in emotional situations. *Neuroimage* **63**, 591–599 (2012).
44. Wittmann, M., van Wassenhove, V., Craig, A. D. B. & Paulus, M. P. The neural substrates of subjective time dilation. *Front. Hum. Neurosci.* **4**, 1–9 (2010).
45. Zakay, D. The roles of non-temporal information processing load and temporal expectations in children's prospective time estimation. *Acta Psychol. (Amst).* **84**, 271–280 (1993).
46. Zakay, D. The roles of non-temporal information processing load and temporal expectations in children's prospective time estimation. *J. Exp. Child Psychol.* **54**, 355–371 (1992).
47. Field, D. T. & Groeger, J. A. Temporal interval production and short-term memory. *Percept. Psychophys.* **66**, 808–19 (2004).
48. Wittmann, M., Simmons, A. N., Aron, J. L. & Paulus, M. P. Accumulation of neural activity in the posterior insula encodes the passage of time. *Neuropsychologia* **48**, 3110–3120 (2010).
49. Deaner, R. O., Khera, A. V. & Platt, M. L. Monkeys pay per view: Adaptive valuation of social images by rhesus macaques. *Curr. Biol.* **15**, 543–548 (2005).
50. Rangel, A., Camerer, C. & Montague, P. R. A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* **9**, 545–556 (2008).
51. Rodriguez, G. Parametric survival models. *Regres. Model. Strateg.* 1–14 at http://link.springer.com/chapter/10.1007/978-1-4757-3462-1_17 (2010).
52. Cox, D. R. Regression Models and Life-Tables. *J. R. Stat. Soc. Ser. B* **34**, 187–220 (1972).
53. Rodriguez, G. Survival Models. *Lecture Notes on Generalized Linear Models* 1–34 at http://data.princeton.edu/wws509/notes/c7.pdf (2010).
54. Prichard, D. & Theiler, J. Generating surrogate data for time series with several simultaneously measured variables. *Phys. Rev. Lett.* **73**, 951–954 (1994).

Acknowledgements

Author Contributions

S.S.C. and L.C.P. designed the experiments, analyzed the data, and wrote the manuscript. S.S.C. and S.H. implemented the experiments. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-04402-4

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.