EEG in the classroom: Synchronised neural recordings during video presentation

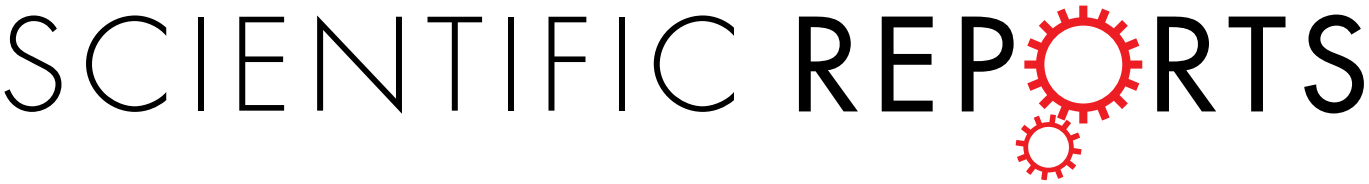
Poulsen, Andreas Trier; Kamronn, Simon Due; Dmochowski, Jacek; Parra, Lucas C.; Hansen, Lars Kai

Published in:
Scientific Reports

Link to article, DOI:
10.1038/srep43916

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record



OPEN

EEG in the classroom: Synchronised neural recordings during video presentation

Received: 26 April 2016
Accepted: 01 February 2017
Published: 07 March 2017

Andreas Trier Poulsen[1,*], Simon Kamronn[1,*], Jacek Dmochowski[2,3], Lucas C. Parra[3] & Lars Kai Hansen[1]

We performed simultaneous recordings of electroencephalography (EEG) from multiple students in a classroom, and measured the inter-subject correlation (ISC) of activity evoked by a common video stimulus. The neural reliability, as quantified by ISC, has been linked to engagement and attentional modulation in earlier studies that used high-grade equipment in laboratory settings. Here we reproduce many of the results from these studies using portable low-cost equipment, focusing on the robustness of using ISC for subjects experiencing naturalistic stimuli. The present data shows that stimulus-evoked neural responses, known to be modulated by attention, can be tracked for groups of students with synchronized EEG acquisition. This is a step towards real-time inference of engagement in the classroom.

Engagement and attention are important in situations of learning, but most methods for measuring of attention or engagement are intrusive and unrealistic in everyday situations[1–3]. Recently, inter-subject correlation (ISC) of electroencephalography (EEG) has been proposed as a marker of attentional engagement[4–6] and we ask in this work whether it can be recorded robustly with commercial-grade wireless EEG devices in a classroom setting. Furthermore, we address two other issues related to the robustness of the signal: The potential neurophysiological origin of the measure and the robustness of the detection scheme to inter-subject variability in spatial alignment.

User engagement has been defined as '… the emotional, cognitive and behavioural connection that exists, at any point in time and possibly over time, between a user and a resource'[7]. Traditional approaches to measuring engagement are based on capturing user behaviour via user interfaces, self-report, or manual annotation[8]. However, tools from cognitive neuroscience are increasingly being employed[9]. Recent efforts in neuroscience aim to elucidate perceptual and cognitive processes in a more realistic setting and using naturalistic stimuli[4,10–14]. From an educational perspective such quantitative measures may help identify mechanisms that make learning more efficient[9], align services better with students needs[7], or monitor critical task performance[15]. The potential uses of engagement detection in the classroom are numerous, e.g., real-time and summary feedback for the teacher, motivational strategies for increased student engagement, and screening for impact of teaching materials. Before the findings of tracking attentional responses with neural activity[4–6] can be employed in a real-time classroom scenario, several issues must be addressed first, including: (1) Is it possible to reproduce the ISCs to naturalistic stimuli under the adverse conditions of a classroom? (2) Are the ISCs robust to inter-student variability of the spatial information processing networks? And (3) can ISCs be recorded with equipment that is both comfortable and affordable enough to make it a realistic technology for schools?

Here we investigate the feasibility of recording such neural responses from students who are viewing videos. We use an approach developed by Dmochowski *et al.*[4] that uses inter-subject correlation (ISC) of EEG evoked responses. The basic premise is that subjects who are engaged with the content exhibit reliable neural responses that are correlated across subjects and repetitions within the same subject. In contrast, a lack of engagement manifests in generally unreliable neural responses[6]. ISC of neural activity while watching films have been shown to predict the popularity and viewership of TV-series and commercials[5], and shows clinical promises as a measure of consciousness levels in non-responsive patients[16] (fMRI study). We argue here that the neural reliability of students indeed may be quantified on a second-by-second basis in groups and in a classroom setting, and we seek

[1]Technical University of Denmark, DTU Compute, Kgs. Lyngby, Denmark. [2]Stanford University, Department of Psychology, Palo Alto, USA. [3]City College of New York, Department of Biomedical Engineering, New York, USA. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to A.T.P. (email: atpo@dtu.dk)

Figure 1. Experimental setup for joint viewings. (**Left**) 9 subjects where placed on a line to induce a cinema-like experiences. (**Right**) Subjects seen from the back, watching films projected onto a screen. Tablets recording EEG are resting on the tables behind the subjects. The signal is transmitted wirelessly from each subject.

to investigate the robustness of measuring it with electroencephalography (EEG) responses during exposure to media stimuli.

To enable correlations between multi-dimensional EEG, correlated component analysis (CorrCA) was introduced[4]. CorrCA finds multiple spatial projections that are shared amongst subjects, such that their components are maximally correlated across time. Here we are interested in the reproducibility of using CorrCA as a measure of inter-subject correlation, and will focus predominantly on the first component, which captures most of the neural responses shared across students.

The main goal of the present work is to determine whether student neural reliability can be quantified in a real-time manner based on recordings of brain activity in a classroom setting using a low-cost, portable EEG system – the Smartphone Brain Scanner[17]. With regard to the robustness of the detection scheme, we report on both theoretical and experimental investigations. First, we show that ISC evoked by rich naturalistic stimuli is robust enough to be reproduced with commercial-grade equipment, and to be recorded simultaneously from multiple subjects in a classroom setting. This opens up for the possibility of real-time estimation of student attentional engagement. Secondly, we show mathematically that the CorrCA algorithm is surprisingly robust to variations in the spatial patterns of brain activity across subjects. Finally, we demonstrate that the level of ISC is related to a very basic visual response that is modulated by narrative coherence of the video stimulus.

Results

To monitor neural reliability we used video stimuli as they provide a balance between realism and reproducibility[11]. We recorded EEG activity using the Smartphone Brain Scanner while subjects watched short video clips of approximately 6 minutes duration, either individually or in a group setting (Fig. 1). To measure reliability of EEG responses, we used correlated components analysis (CorrCA, see Methods) to extract maximally correlated time series with shared spatial projection across repeated views within the same subject (inter-viewing correlation, IVC), or between subjects (inter-subject correlation, ISC).

One of our main points of interest is to investigate the robustness of ISC from EEG recorded in a classroom through comparisons with results previously measured in a laboratory setting[4]. We therefore employed similar methods of analysis and calculated ISCs and IVCs in 5 second windows with 80% overlap to investigate their temporal development in a 1-second resolution. We chose to analyse the EEG with CorrCA in a broad frequency band (0.5 and 45 Hz), instead of investigating specific frequency bands, to keep the analysis methods comparable with the prior lab-based study. Moreover, CorrCA is a method used for robustly measuring ISC with low computational costs; hence making it a good candidate for long term real-time analyses on small devices in a classroom setting.

The subjects watched three video clips, which were presented twice in random order. The first video was a suspenseful excerpt from the short film, *Bang! You're Dead*, directed by Alfred Hitchcock. It was selected because it is known to effectively synchronize brain responses across viewers[4,18]. The second video was an excerpt from *Sophie's Choice*, directed by Alan J. Pakula (1982), and the third was an uneventful baseline video of people silently descending an escalator. For both the joint and individual recording scenarios, the time course of the ISC, based on the first CorrCA component from subjects watching the film, closely reproduces results obtained previously in a laboratory setting (Fig. 2a and Table 1).

An indication of the stability of the technique is provided by the spatial patterns of the neural activity that drives these reproducible responses. Similar to other component extraction techniques, such as independent component analysis or common spatial patterns[19,20], CorrCA reduces the signal of multiple electrodes to a few components. The ISC is then computed for the first few components, which capture most of the correlation between recordings. The strongest three correlated components show a stable pattern of activity across the different groups and recording conditions (Fig. 2b), all three obtaining significant spatial correlations between groups ($r_{comp1} = 0.97$, $r_{comp2} = 0.91$, $r_{comp3} = 0.79$, all with $p < 0.002$ for uncorrected permutation test), for *Bang! You're Dead*. The robustness to recording conditions is also apparent for the second film clip from *Sophie's Choice* ($r_{comp1} = 0.51$, $p < 0.002$; $r_{comp2} = 0.48$, $p = 0.008$; $r_{comp3} = 0.36$, $p = 0.033$), albeit with a lower average correlation, which for the first two components may be due to noisy scalp maps for the *Joint 1* group and *Individual* group,

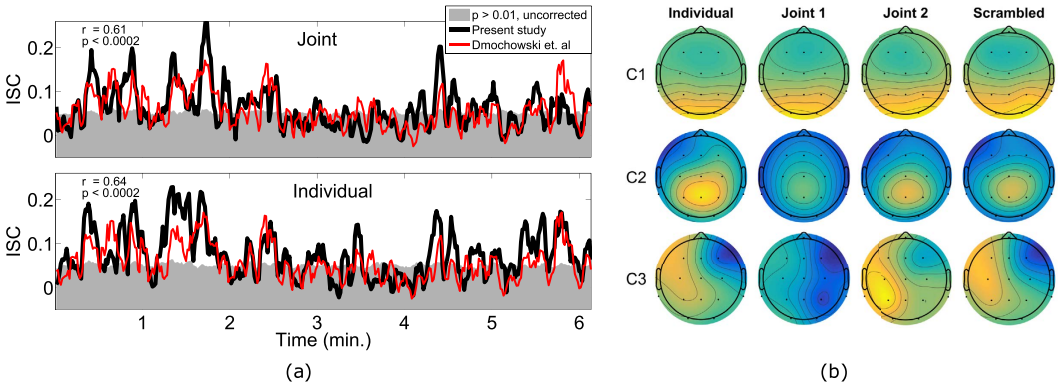


Figure 2. ISC of neural responses to naturalistic stimuli are robust across different groups of subjects and reproducible in a classroom setting. (**a**) Comparison between the ISC obtained by Dmochowski *et al.*[4] and the present study for the first CorrCA component and the first viewing of *Bang! You're Dead*. The ISC is calculated with a 1-second resolution (5 s windows, 80% overlap). The grey area indicates chance levels for ISC ($p > 0.01$ estimated with time-shuffled surrogate data, uncorrected for multiple comparisons). (**b**) The corresponding scalp projections of the first three components obtained from the correlated component analysis (CorrCA) of each of the four subject groups watching *Bang! You're Dead* the first time. For each component, CorrCA finds one shared set of weights for all subjects in the group. Four distinct groups of subjects watched videos in different scenarios: individually on a tablet computer (*Individual*), individually with order of scenes scrambled in time (*Scrambled*), and jointly in a classroom as seen in Fig. 1 (*Joint 1* and *Joint 2*). For each projection, the polarity was normalized so the value at the Cz electrode is positive.

	ISC v1	ISC v2	IVC
Individual	0.64**	0.33**	0.49**
Joint group 1	0.51**	0.15**	0.44**
Joint group 2	0.61**	0.28**	0.54**

Table 1. Correlation coefficients between the ISC time courses obtained in a laboratory setting[4] and those obtained in the present study (groups *Individual, Joint 1* and *Joint 2*). Inter-subject correlation (ISC) measures similarity of responses between subjects for the first and second viewings (v1, v2), and the inter-viewing correlation (IVC) measures similarity within-subject between the two views. Coefficients are calculated for the first CorrCA component recorded while watching *Bang! You're dead*. **$p < 0.01$.

respectively (see Supplementary Fig. S1). For the baseline video, only the first component achieved significant average correlation between groups ($r_{comp1} = 0.46$, $p = 0.014$). The lower stability in the scalp maps obtained for *Sophie's Choice* and the baseline video could be explained by the lower ALD of these stimuli (see below), since these films obtain lower average IVC compared to *Bang! You're Dead* for all groups (Fig. 3).

Previous research has indicated the potentials of ISC as a marker of engagement of conscious processing[4,5,6,12,16]. To further investigate this, we asked subjects post-experiment to describe the film segments (or "scenes") that made the biggest impact on them. We quantified their answers by assigning each answer to one of eight general scene descriptions. Table 2 shows that the scenes most frequently mentioned are "Boy pointing gun at mother" or "Boy pointing gun at people", and 29 out of 30 subjects mentioned one or both of the scenes as having had high impact on them. The most frequently mentioned scene occurs around 2:25, where a peak in the ISC can be seen (Fig. 2a). The high impact of this particular scene was confirmed by the suspense ratings presented in Naci *et al.*[16]. See Dmochowski *et al.*[4] for additional descriptions and examples of scenes eliciting high ISC in *Bang! You're Dead*.

To determine if the portable equipment, which uses only 14 channels, can detect varying levels of neural reliability, a second group of subjects watched the same two film clips individually, but now with scenes scrambled in time. This intervention is a widely used tool to create a baseline with similar low-level stimuli, yet reduced engagement[4,18,21,22]. See Methods for more information on the definition and time scales of the scrambled scenes. Despite using consumer-grade EEG we find that IVC is significantly above chance for a large fraction of the original engaging clip, but drops dramatically when the scenes are scrambled in time (mean IVC, Fig. 3, $p < 0.01$, for *Bang! You're Dead*). Also the baseline video, which subjects reported not to engage them at all, only obtained significant ISC ($p < 0.01$, uncorrected) in 2.3% of the 354 tested time windows, compared to the 54.1% significant windows obtained for *Bang! You're Dead*.

For experiments conducted in less controlled, everyday settings as in this study, it is important to assess across-session reproducibility. To test this, we recorded a second group of subjects in a classroom setting who watched the material together (*Joint 1* and *2*). These two groups obtained mean IVCs comparable to the individual recordings (Fig. 3, *Bang! You're Dead*: $p > 0.49$, *Sophie's Choice*: $p > 0.26$), and also showed reproducibility between the groups of simultaneous recordings (Fig. 3, *Bang! You're Dead*: $p > 0.49$, *Sophie's Choice*: $p > 0.08$).

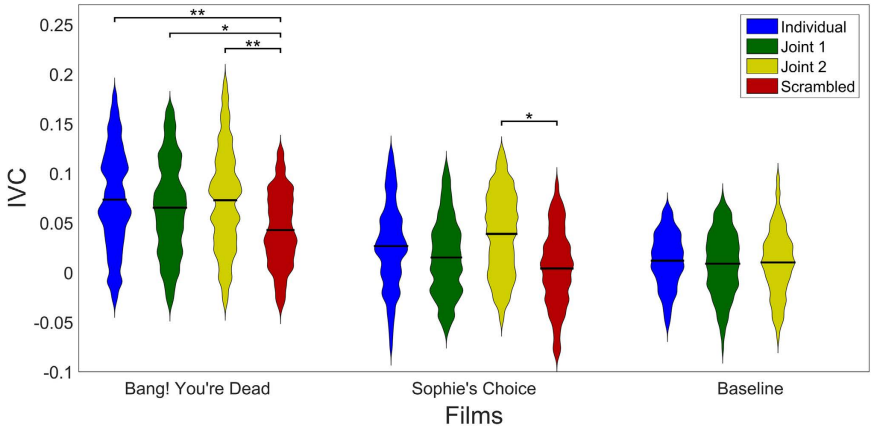


Figure 3. Distribution and mean of IVC calculated from the first CorrCA component for subject groups and films. Violin plots show distributions of IVC estimated using a squared exponential (normal) kernel with bandwidth of 0.005[41]. Horizontal black bars denote distribution means. For visualisation purposes, the extreme 2.5% values at either end of the distributions were left out of the violin plots (but were kept for estimating mean and p-values). A block permutation test (block size $B = 25\,s$) was employed to estimate statistical significant differences in the mean IVC between viewing conditions (uncorrected for multiple comparisons). For both films there were significant differences in mean IVC between groups with normal narrative and the *Scrambled* group (*Bang! You're Dead*: $p_{Individual} = 0.006$, $p_{Joint1} = 0.033$, $p_{Joint2} = 0.004$; *Sophie's Choice*: $p_{Individual} = 0.059$, $p_{Joint1} = 0.37$, $p_{Joint2} = 0.012$). However, there were no significant differences between groups with the original, unscrambled narrative. Note that the *Scrambled* group did not watch the baseline video.

Scene	Approx. times	No of times mentioned (%)
The boy shoots (or points gun at) mother	2:25 and 3:00	16 (53%)
The boy shoots (or points gun at) at people	2:10, 3:30 and 5:30	15 (50%)
The boy loads another bullet into gun	6:10	8 (27%)
The uncle discovers his gun is gone	4:35	4 (13%)
The boy finds and loads gun	0:25 and 1:40	4 (13%)
The boy points at mirror or shoot towards camera	0:40, 1:50 and 5:25	4 (13%)
When the father did not run after the boy	3:00	1 (3%)
The abrupt ending	6:14	1 (3%)

Table 2. Scenes described by the subjects as having the strongest impression on them. Based on the 30 subjects which saw *Bang! You're Dead* with uninterrupted narrative. In a post-experiment questionnaire, subjects were asked to describe the scenes that made the strongest impression on them. Their answers were collected in the eight groups. The subjects each mentioned 1.77 scenes on average (0.77 std.). 29 subjects (97%) mentioned either scenes where the boy points the gun at his mother or at other people.

Robustness to inter-subject variations in the spatial brain structure is a basic question when applying CorrCA to classroom data. CorrCA is derived under the assumption that the spatial networks of subjects are identical. This assumption could be challenged by inter-individual differences, however, it turns out to be surprisingly robust to such variability[23]. To demonstrate this, we briefly analyse a 'worst case' scenario in which the true mixing weights of two subjects form a pair of *orthogonal* vectors. The observations are assumed to consist of a single true signal, $\mathbf{z}$, mixed into D dimensions with additive Gaussian noise; $\mathbf{X}_1 = \mathbf{a}_1\mathbf{z}^\top + \varepsilon$, $\mathbf{X}_2 = \mathbf{a}_2\mathbf{z}^\top + \varepsilon$. Given a large sample, the covariance matrices are given as $\mathbf{R}_{11} = P \cdot \mathbf{a}_1\mathbf{a}_1^\top + \sigma^2\mathbf{I}$, $\mathbf{R}_{12} = P \cdot \mathbf{a}_1\mathbf{a}_2^\top$, where P is the variance of $\mathbf{z}$ and σ^2 signifies the noise variance. For simplicity the weight vectors are assumed to be unit length. The two matrices in Eq. (3) can then be written as

$$(\mathbf{R}_{11} + \mathbf{R}_{22})^{-1} = \frac{1}{P}\left([\mathbf{a}_1\,\mathbf{a}_2]\begin{bmatrix}\mathbf{a}_1^\top \\ \mathbf{a}_2^\top\end{bmatrix} + \frac{2\sigma^2}{P}\mathbf{I}\right)^{-1};\quad \mathbf{R}_{12} + \mathbf{R}_{21} = P \cdot [\mathbf{a}_1\,\mathbf{a}_2]\begin{bmatrix}\mathbf{a}_2^\top \\ \mathbf{a}_1^\top\end{bmatrix}, \tag{1}$$

using block matrix notation. With $\mathbf{a}_1^\top\mathbf{a}_2 = 0, \|\mathbf{a}_1\|^2 = \|\mathbf{a}_2\|^2 = 1$ and the Woodbury identity, the product of the two matrices in Eq. (1) can be expressed as

$$(\mathbf{R}_{11} + \mathbf{R}_{22})^{-1}(\mathbf{R}_{12} + \mathbf{R}_{21}) = \frac{P}{2\sigma^2 + P}(\mathbf{a}_1\mathbf{a}_2^\top + \mathbf{a}_2\mathbf{a}_1^\top). \tag{2}$$

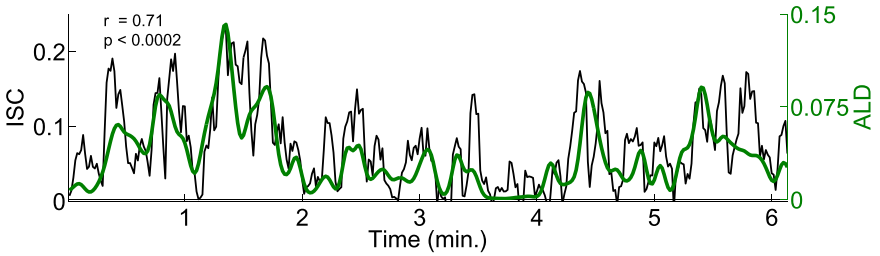


Figure 4. The ISC of the first CorrCA component is temporally correlated with the average luminance differences (ALD) of the film stimulus. ALD is calculated as the frame-to-frame difference in pixel intensity, smoothed to match the 5 s window of ISC, and mainly reflects the frequency of changes in camera position. Data computed from the neural responses of subjects watching *Bang You're Dead*.

	ISC v1	ISC v2	IVC
Bang You're Dead	0.71**	0.61**	0.56**
Sophie's Choice	0.50**	0.24**	0.23**
Bang You're Dead (Scr)	0.54**	0.45**	0.35**
Sophie's Choice (Scr)	0.42**	0.01	−0.22**

Table 3. Correlation coefficients between the ALD and the ISC for the two viewings (v1, v2) as well as the IVC for the first correlated component. The correlation is presented for *Bang You're dead* and *Sophie's Choice* for the *Individual* and *Scrambled (Scr)* groups. $^{**}p < 0.01$.

An eigenvector of matrix (2) takes the form $\alpha\mathbf{a}_1 + \beta\mathbf{a}_2$, with $\alpha = \pm\beta$ and $\pm\frac{P}{2\sigma^2 + P}$ as eigenvalues. By applying this eigenvector to observations, $\mathbf{X}_1$ and $\mathbf{X}_2$, we see that CorrCA still identifies the relevant time series, $\mathbf{z}$.

For the first CorrCA component, the channels weighted most heavily are the ones positioned over the occipital lobe (see Fig. 2b). To estimate how much of the ISC was driven by basic low-level visual processing, we analysed the relation between ISC and a measure of frame-to-frame luminance fluctuations (average luminance difference, ALD; see methods). Note that to avoid synchronised eye artefacts and to ensure that only signals of neural origin contributed to the measured correlations, we removed independent components related to eye artefacts from the EEG (see methods).

Figure 4 and Table 3 show that there is a significant correlation between the ISC and the ALD for both *Bang! You're Dead* and *Sophie's Choice* for the first CorrCA component. This suggests that this portion of the correlated activity may indeed be driven by low-level visual evoked responses. However, the degree of engagement, here represented by narrative coherence, appears to modulate the *amplitude* of the ISC time course, since even though the scrambled stimulus was driven by the visual stimulus, it was so to a lesser extent. Previous research has shown that visual evoked potentials (VEP) are modulated by spatial attention[24] and that even feature-specific attention enhances steady-state VEPs[25]. We quantify the effect of scrambling the narrative by comparing the sensitivity (slope) of ISC to ALD in both the normal and scrambled conditions by fitting a simple linear model (Fig. 5). For both films we found significant reductions of the ISC/ALD slope in the scrambled version ($p < 0.01$; block permutation test, with block size $B = 25\,s$).

Discussion

We have demonstrated that student neural reliability to media stimuli may be quantified using EEG in a classroom setting. For educational technology cost and robustness are key features, hence, we aimed at establishing a realistic scenario based on low-cost consumer grade equipment, the Smartphone Brain Scanner, focusing on several potential sources that could degrade robustness.

We have provided evidence that salient aspects of the neural reliability previously detected with laboratory grade equipment can be reproduced in a realistic setting. We recorded fully-synchronized EEG with nine subjects in a real classroom and found that the level of neural response reliability matched prior laboratory results. The robustness of CorrCA and ISC is granted by the reproducibility between recording conditions, both of the ISC time-courses throughout the film clips and of the spatial topographies of the first three CorrCA components. For the film clip from *Bang! You're Dead* we saw that seven subjects were enough to obtain stable topographies for all three components, whereas for *Sophie's Choice* and the baseline video the results were more noisy, suggesting that more subjects are needed to obtain stable results. Previous research shows that ten subjects provided for stable results in a case involving non-narrative baseline videos or films with lower ISC and IVC in a laboratory setting[4].

Mathematically, we have shown that our detection scheme, CorrCA, is robust to inter-subject variability in spatial configurations of brain networks, or induced by cap misalignment. In the calculations, we assumed two subjects in a worst case scenario where the subjects' spatial projections are orthogonal. This result conforms well with simulations that show that, even for multiple subjects with randomly drawn spatial projections, CorrCA was able to find the relevant times series[23]. The simulations also showed that increasing the number of subjects

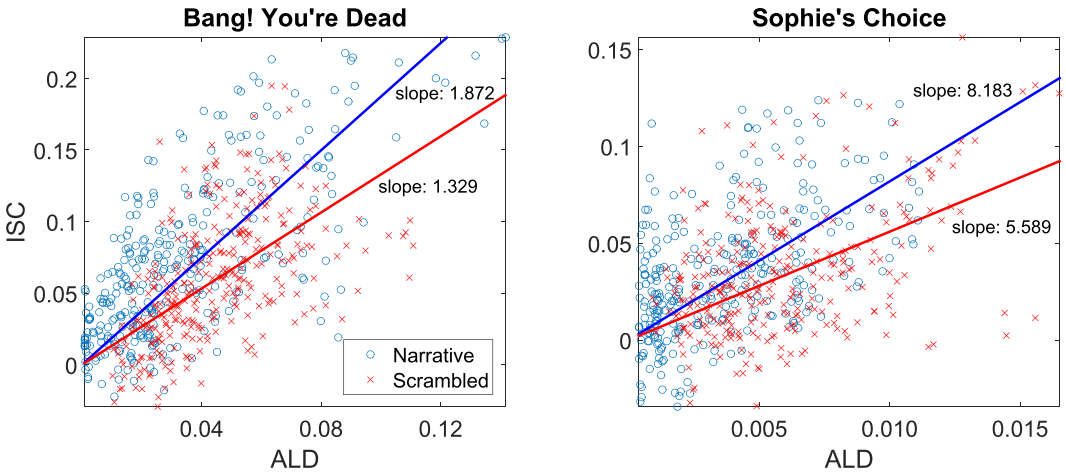


Figure 5. Relation between the ISC and the ALD for different conditions. Each point indicates a point in the ISC time course as seen in Fig. 2a (5 s windows, 80% overlap) and the corresponding ALD calculated from the visual stimulus. It is evident that time points with higher luminance fluctuations (hight ALD) result in higher correlation of brain activity across subjects (high ISC). The indicated "slope" is a least squares fit of the slope of lines passing through (0, 0). The slope indicates the strength of ISC for a given ALD value. For both films there is a significant drop in the slope ($p < 0.01$: block permutation test with block size $B = 25$ sec), thus the original narrative (blue) elicits higher ISC than the less engaging scrambled version of the films (red). Note that brightness of the scenes in *Sophie's Choice* is much lower than in *Bang! You're dead*, resulting in an ALD that is lower by almost a factor 10.

decreased the signal-to-noise ratio, presumably due to the estimated common projection not being able to fit with the different projections of each subject.

We have presented results that further indicate a relationship between changes in ISC and viewer engagement. Through a basic analysis of questionnaires on scenes of high impact, we found that high ISC indeed is associated with high impact. We have also showed a relationship between neural responses to luminance fluctuations and coherence of stimulus narrative. For both the films presented, we saw a significant drop in the average IVC for subjects watching the film sequences in which the narrative had been temporally scrambled. At the same time no significant difference was found between the groups watching the film sequences that had not been scrambled, which further underlines the robustness of the measure.

It may appear surprising that there exists a significant correlation between the *raw EEG signals* of various students in the classroom. However, it is well-known that eye scan patterns in a film audience follow a specific pattern after a scene change, activating the dorsal pathway[26]. A valid assumption could therefore be that the correlation is due to synchronised artefacts from eye movements, but this has recently been shown not to affect attentional modulation of ISC[6]. Also, it is known that stimuli in the form of flashing images elicit VEPs, which are modulated in amplitude by the luminance[27]. When recorded with EEG, the spatial distribution of the early VEP at 100 ms (P100) is similar to the scalp maps of the first correlated component (C1 in Fig. 2b)[24,28].

We investigated whether low-level visual processes could be a driving force behind the measured ISCs by correlating the ISC with changes in luminance in the video stimuli, as measured by the ALD. We found that luminance fluctuations drive a significant portion of the ISC.

In all four groups of subjects *Sophie's Choice* obtained lower IVC compared to *Bang! You're Dead*. This difference could be explained by the fact that the film clip also had a much lower ALD. Also, Fig. 4 indicates that the passage in *Bang! You're dead* with the highest and most sustained ISC (around 1:20 to 1:50) coincides with the interval with the most scene changes. This relationship could, however, also be due to more complex processes, as fast-paced cutting is a known cinematographic tool used by Hitchcock to induce suspense and thereby increase the attention of the viewer[29].

The strong link between ISC and luminance fluctuations due to scene cuts has also recently been presented in a fMRI study[30]. This is something that would be interesting to take into account for future studies investigating the applicability of ISC. Baseline videos could be created in ways to achieve similar ALD features as the target stimuli. The baseline video, created for this study, consisted of one continuous scene of people entering and exiting an escalator in a relaxed manner, which did not produce any significant correlation. Future studies might use a baseline video containing scene cuts of faces and body parts, to also take the effect of editing into account.

To investigate the possibility of higher level processes also being at play, we analysed the linear relationship between ISC and luminance fluctuations at a given time in the video stimulus. The scrambling operation aimed to test for a change in attentional engagement while controlling for low level features. The premise was that subjects would be less attentive to the stimulus, i.e. less "engaged", if they did not follow the narrative arch of the story. With that in mind, Figs 4 and 5 suggest that ISC is driven by stimulus-evoked responses that are modulated by attentional engagement with the stimulus.

We have demonstrated the feasibility of tracking inter-subject correlation in a classroom setting; a measure that has been related to attentional modulation[6]. We have shown that ISC is robust to recording equipment and

conditions, and we have presented evidence that the amplification of ISC in films that have a strong and coherent narrative is due to attentional modulation of visual evoked responses. Thus ISC may be used as an indirect electrophysiological measure of engagement through an attentional top-down modulation of low-level neural processes. Recent research has shown that attentional modulation of neural responses takes place in speech perception[31,32], which lends credibility to a similar process occurring in the visual system. The evidence that such a basic and well defined mechanism could be at play further adds to the robustness of the approach in real everyday scenarios.

Methods

Protocol. Four groups of subjects watched the video stimuli in different scenarios. The first group ($N = 12$, *Individual*) watched videos individually in an office environment on a tablet computer (Google Nexus 7 tablet, with a 7″ (17.8 cm) screen) with earphones. The second group ($N = 12$) saw the videos in the same manner, but the scenes of the film stimulus were scrambled in time resulting in the narrative being lost (*Scrambled*). The objective of this condition was to demonstrate that the similarity of responses across subjects is not simply the result of low-level stimulus features (which are identical in the *Individual* and *Scrambled* conditions), but instead, is modulated narrative coherence, which presumably engages viewers. Two additional groups ($N = 9$, $N = 9$) watched the original videos on a screen in a classroom (Fig. 1, *Joint 1* and *Joint 2*), with sound projected through loudspeakers. An attempt was made to create viewing conditions for the subjects in the *joint* groups, that were similar to the viewing conditions for the *individual* group, i.e., lights were dampened and the projected image produced approximately the same field-of-view (see Supplementary materials). The central question was whether the viewing condition (i.e., in a group versus individually) influences the level of ISC across subjects.

Stimuli. The first video clip was a suspenseful excerpt from the short film *Bang! You're Dead* (1961) directed by Alfred Hitchcock. It was selected because it is known to elicit highly reliable brain activity across subjects in fMRI[11] as well as EEG[4]. Our second stimulus was a clip from *Sophie's Choice*, directed by Alan J. Pakula (1982), which has been used earlier to study fMRI activity in the context of emotionally salient naturalistic stimuli[33]. A third non-narrative control video was recorded in a Danish metro station of several people who were being transported quietly on an escalator. Each video clip had a length of approximately six minutes and was shown twice to each subject. For each viewing the order of the clips was randomized, while the same random order was used the second time the clips were shown. A combined video was created for each of the six possible permutations of the order of the clips, starting with a 10 second 43 Hz tone for use in post processing synchronization, and 20 seconds black screen between each film clip. The total length of the video amounted to 39 minutes. An additional control stimulus (*Scrambled*) was created by scrambling the order of the scenes in *Bang! You're Dead* and *Sophie's Choice* in accordance with previous research[4,18]. In these studies, scene segments were defined in varying temporal scales (36 s, 12 s, and 4 s) that consisted of multiple camera positions, "shots". For this study we defined a scene as a single shot (i.e. the segment between two scene cuts) with the added rule that a scene must not exceed 250 frames (~10 s) to reduce subjects' ability to infer the narrative from long scenes. This procedure resulted in 73 scenes lasting between 0.5 and 10 seconds and corresponded to the intermediate to short time-scales employed in previous studies[18].

Subjects. A total of 42 female subjects (mean age: 22.4 y, age range: 18–32 y), who gave written informed consent prior to the experiment, were recruited for this study. Non-invasive experiments on healthy subjects are exempt from ethical committee processing by Danish law[34]. Among the 42 recordings, nine were excluded due to unstable wireless communication that precluded proper synchronization of the data across subjects (five from the *Individual* group, one from the *Scrambled* group and three from the two *Joint* groups). The difference in the number of recordings in the different groups could give unfair advantages with respect to noise when using CorrCA or calculating ISC. We therefore decided to randomly choose four subjects from the *Scrambled* group and one from *Joint 2* group and excluded these from the analyses. This was to ensure that each group had seven fully synchronized recordings.

Portable EEG – Smartphone Brain Scanner. Research grade EEG equipment is costly, time-consuming to set up, and immobile. However, recently consumer grade EEG equipment that is more affordable and has increased comfort has appeared. Here we use the modified 14 channel system, 'Emocap', based on the EEG Emotiv EPOC headset. For details and validation, see refs 17 and 35. In this study it was implemented on Asus Nexus 7 tablets. An electrical trigger and associated sound was used to synchronize EEG and video signals in the individual viewing condition, while a split audio signal (simultaneously feeding into microphone and EEG amplifiers) was used to synchronize the nine subjects EEG recordings and the video in the joint viewing condition (see Supplementary materials for further information on synchronisation). The resulting timing uncertainty was measured to be less than 16 ms. The EEG was recorded at 128 Hz and subsequently bandpass filtered digitally using a linear phase windowed sinc FIR filter between 0.5 and 45 Hz and shifted to adjust for group delay. Eye artefacts were reduced with a conservative pre-processing procedure using independent component analysis (ICA), removing up to 3 of the 14 available components (Corrmap plug-in for EEGLAB[36,37]).

Correlated component analysis to measure ISC and IVC. CorrCA was presented in Dmochowski *et al.*[4], as a constrained version of Canonical Correlation Analysis (CCA). CorrCA seeks to find sets of weights that maximises the correlation between the neural activity of subjects experiencing the same stimuli. For each neural component, CorrCA finds one shared set of weights for all subjects in the group.

Given two multivariate spatio-temporal time series (termed "view" in CorrCA), $\{\mathbf{X}_1, \mathbf{X}_2\} \in \mathbb{R}^{D\times N}$, with D being the number of measured features (EEG channels) in the two views and N the number of time samples,

CCA estimates weights, $\{\mathbf{w}_1, \mathbf{w}_2\}$, which maximize the correlation between the components, $\mathbf{y}_1 = \mathbf{X}_1^\top \mathbf{w}_1$ and $\mathbf{y}_2 = \mathbf{X}_2^\top \mathbf{w}_2$. The weights are calculated using two eigenvalue equations, with the constraint that the components belonging to each multivariate time series are uncorrelated[38]. CorrCA is relevant for the case where the views are homogeneous, e.g., using the same EEG channel positions, and imposes the additional constraint of shared weights $\mathbf{w} = \mathbf{w}_1 = \mathbf{w}_2$. This assumption can potentially increase sensitivity involving fewer parameters. In CorrCA the weights are thus estimated through a single eigenvalue problem;

$$(\mathbf{R}_{11} + \mathbf{R}_{22})^{-1}(\mathbf{R}_{12} + \mathbf{R}_{21})\mathbf{w} = \rho\mathbf{w}, \tag{3}$$

where, $\mathbf{R}_{ij} = \frac{1}{N}\mathbf{X}_i\mathbf{X}_j^\top$, is the sample covariance matrix[4]. To illustrate the spatial distribution of the underlying physiological activity of the components, we use the estimated forward models ("patterns") as discussed in refs 39 and 40.

Average luminance difference (ALD). Video clips were converted to grey scale (0–255) by averaging over the three colour channels. We then calculated the squared difference in pixel intensity from one frame to the next and took the average across pixels. These signals were non-linearly re-sampled at 1 Hz by selecting the maximum ALD for each 1 *s* interval to emphasise the large differences during changes in camera position (see Figure S2 in Supplementary materials for an comparison between frame-to-frame and smoothed difference). These values were then smoothed in time by convolving with a Gaussian kernel with a "variance" parameter of $2.5\,s^2$. This down sampling and smoothing was aimed at matching the temporal resolution of the ALD to that of the time-resolved ISC computation (5 *s* sliding window with 1 *s* intervals).

Statistical testing. In order to evaluate the statistical relevance of the correlations, we employed a simple permutation test ($P = 5000$ permutations)[4]. To test the robustness of the obtained weights for the spatial projections, we calculated the average correlation of all possible pairings of the four conditions groups for a given component. Again, we employed a permutation test ($P = 5000$ permutations) to evaluate statistical relevance by randomly permuting the channel order for each group and recalculating the average correlation. When testing differences in average IVC between conditions, we used a block permutation test (block size $B = 25\,s$, $P = 5000$ permutations) to account for temporal dependencies.

References

1. Robinson, P. Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning* **47,** 45–99 (1997).
2. Cohen, A., Ivry, R. B. & Keele, S. W. Attention and structure in sequence learning. *J. Exp. Psychol.* [*Learn. Mem. Cogn.*] **16,** 17–30 (1990).
3. Radwan, A. A. The effectiveness of explicit attention to form in language learning. *System* **33,** 69–87 (2005).
4. 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? *Frontiers in human neuroscience* **6,** 112 (2012).
5. Dmochowski, J. P. *et al.* Audience preferences are predicted by temporal reliability of neural processing. *Nature Communications* **5,** 1–9 (2014).
6. Ki, J. J., Kelly, S. P. & Parra, L. C. Attention strongly modulates reliability of neural responses to naturalistic narrative stimuli. *The Journal of Neuroscience* **36,** 3092–3101 (2016).
7. Attfield, S., Kazai, G., Lalmas, M. & Piwowarski, B. Towards a science of user engagement. In *WSDM Workshop on User Modelling for Web Applications* (ACM International Conference on Web Search And Data Mining, 2011).
8. O'Brien, H. L. & Toms, E. G. Examining the generalizability of the user engagement scale (ues) in exploratory search. *Information Processing & Management* **49,** 1092–1107 (2013).
9. Szafir, D. & Mutlu, B. Artful: adaptive review technology for flipped learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1001–1010 (ACM, 2013).
10. Ringach, D. L., Hawken, M. J. & Shapley, R. Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences. *Journal of vision* **2,** 2–2 (2002).
11. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. Intersubject synchronization of cortical activity during natural vision. *science* **303,** 1634–1640 (2004).
12. Lahnakoski, J. M. *et al.* Synchronous brain activity across individuals underlies shared psychological perspectives. *NeuroImage* **100,** 316–24 (2014).
13. Lankinen, K., Saari, J., Hari, R. & Koskinen, M. Intersubject consistency of cortical MEG signals during movie viewing. *NeuroImage* **92,** 217–224 (2014).
14. Chang, W.-T. *et al.* Combined MEG and EEG show reliable patterns of electromagnetic brain activity during natural viewing. *NeuroImage* **114,** 49–56 (2015).
15. Lin, C.-T., Huang, K.-C., Chuang, C.-H., Ko, L.-W. & Jung, T.-P. Can arousing feedback rectify lapses in driving? Prediction from eeg power spectra. *Journal of neural engineering* **10,** 056024 (2013).
16. Naci, L., Sinai, L. & Owen, A. M. Detecting and interpreting conscious experiences in behaviorally non-responsive patients. *NeuroImage* (2015).
17. Stopczynski, A., Stahlhut, C., Larsen, J. E., Petersen, M. K. & Hansen, L. K. The Smartphone Brain Scanner: A Portable Real-Time Neuroimaging System. *PloS one* **9,** e86733 (2014).
18. Hasson, U. *et al.* Neurocinematics: The neuroscience of film. *Projections* **2,** 1–26 (2008).
19. Parra, L. & Sajda, P. Blind source separation via generalized eigenvalue decomposition. *The Journal of Machine Learning Research* **4,** 1261–1269 (2003).
20. Koles, Z. J., Lazar, M. S. & Zhou, S. Z. Spatial patterns underlying population differences in the background eeg. *Brain topography* **2,** 275–284 (1990).
21. Miller, G. A. & Selfridge, J. A. Verbal context and the recall of meaningful material. *The American journal of psychology* **63,** 176–185 (1950).
22. Anderson, D. R., Fite, K. V., Petrovich, N. & Hirsch, J. Cortical activation while watching video montage: An fmri study. *Media Psychology* **8,** 7–24 (2006).
23. Kamronn, S., Poulsen, A. T. & Hansen, L. K. Multiview Bayesian Correlated Component Analysis. *Neural Computation* **27,** 2207–2230 (2015).

24. Johannes, S., Münte, T. F., Heinze, H. J. & Mangun, G. R. Luminance and spatial attention effects on early visual processing. *Cognitive Brain Research* **2,** 189–205 (1995).
25. Müller, M. M. *et al.* Feature-selective attention enhances color signals in early visual areas of the human brain. *Proceedings of the National Academy of Sciences of the United States of America* **103,** 14250–4 (2006).
26. Unema, P. J., Pannasch, S., Joos, M. & Velichkovsky, B. M. Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition* **12,** 473–494 (2005).
27. Armington, J. C. The electroretinogram, the visual evoked potential, and the area-luminance relation. *Vision research* **8,** 263–276 (1968).
28. Sandmann, P. *et al.* Visual activation of auditory cortex reflects maladaptive plasticity in cochlear implant users. *Brain* **135,** 555–568 (2012).
29. Bordwell, D. Intensified continuity: visual style in contemporary American film. *Film Quarterly* **55,** 16–28 (2002).
30. Herbec, A., Kauppi, J. P., Jola, C., Tohka, J. & Pollick, F. E. Differences in fMRI intersubject correlation while viewing unedited and edited videos of dance performance. *Cortex* **71,** 341–348 (2015).
31. Mesgarani, N. & Chang, E. F. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* **485,** 233–236 (2012).
32. Mirkovic, B., Debener, S., Jaeger, M. & De Vos, M. Decoding the attended speech stream with multi-channel EEG: implications for online, daily-life applications. *Journal of Neural Engineering* **12,** 046007 (2015).
33. Raz, G. *et al.* Portraying emotions at their unfolding: a multilayered approach for probing dynamics of neural networks. *NeuroImage* **60,** 1448–61 (2012).
34. Den Nationale Videnskabsetiske Komité. Vejledning om anmeldelse, indberetning mv. (sundhedsvidenskablige forskningsprojekter) (2014).
35. Stopczynski, A. *et al.* Smartphones as pocketable labs: visions for mobile brain imaging and neurofeedback. *International journal of psychophysiology* **91,** 54–66 (2014).
36. Delorme, A. & Makeig, S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods* **134,** 9–21 (2004).
37. Viola, F. C. *et al.* Semi-automatic identification of independent components representing EEG artifact. *Clinical Neurophysiology* **120,** 868–877 (2009).
38. Hardoon, D. R., Szedmak, S. & Shawe-taylor, J. Canonical correlation analysis; An overview with application to learning methods. *Neural computation* **16,** 2639–2664 (2004).
39. Parra, L. C., Spence, C. D., Gerson, A. D. & Sajda, P. Recipes for the linear analysis of EEG. *NeuroImage* **28,** 326–41 (2005).
40. Haufe, S. *et al.* On the interpretation of weight vectors of linear models in multivariate neuroimaging. *NeuroImage* **87,** 96–110 (2014).
41. Hoffmann, H. violin. m-Simple violin plot using matlab default kernel density estimation (2015).

Acknowledgements

We thank the reviewers for their many constructive comments that greatly improved the manuscript. We thank Ivana Konvalinka, Arek Stopczynski and the DTU Smartphone Brain Scanner team for their assistance and helpful discussions. This work was supported by the Lundbeck Foundation through the *Center for Integrated Molecular Brain Imaging* and by Innovation Foundation Denmark through *Neurotechnology for 24/7 brain state monitoring*.

Author Contributions

A.T.P., S.K., J.D., L.P. and L.K.H. designed research; S.K., A.T.P. and L.K.H. performed research; A.T.P., S.K., J.D., L.P., and L.K.H. contributed analytical tools; A.T.P., S.K., L.K.H. analysed data; A.T.P., S.K., J.D., L.P., and L.K.H. wrote the paper.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing Interests: The authors declare no competing financial interests.

How to cite this article: Poulsen, A. T. *et al.* EEG in the classroom: Synchronised neural recordings during video presentation. *Sci. Rep.* **7**, 43916; doi: 10.1038/srep43916 (2017).

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