Measuring engagement in a classroom: Synchronised neural recordings during a video presentation

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

[1]Technical University of Denmark, DTU Compute, Kgs. Lyngby, Denmark
[2]Stanford University, Department of Psychology, Pala Alto, USA
[3]City College of New York, Department of Biomedical Engineering, New York, USA
[*]lkai@dtu.dk
[+]These authors contributed equally to this work

April 12, 2016

Abstract

The efficacy of learning in a classroom depends on how engaged students are with the learning material. Is it possible to assess students' engagement, not with questionnaires or surveys, but directly from their brain activity without distracting them from the material at hand? We demonstrate that classroom engagement can be measured using students' neural responses to a naturalistic stimulus. Many aspects of attentional modulation, which were previously demonstrated in a laboratory setting are reliably reproduced here in a classroom with portable low-cost electroencephalography (EEG) recorded simultaneously from multiple students. The present data suggests that evoked neural responses to video stimulus, known to be modulated by attention, can be used to assess the engagement of a group of students in real-time.

Introduction

We are interested in quantifying student engagement with digital media in a classroom setting. Measuring student engagement is of importance to both neuroscience and educational research (Attfield et al., 2011; Dmochowski et al., 2012; Lahnakoski et al., 2014; Dmochowski et al., 2014). Recent efforts in neuroscience aim to elucidate perceptual and cognitive processes in a more realistic setting and using naturalistic stimuli (Ringach et al., 2002; Hasson et al., 2004; Dmochowski et al., 2012). From an educational perspective we may identify mechanisms that make learning more efficient (Szafir and Mutlu, 2013), align services better with students needs (Attfield et al., 2011), or monitor critical task performance (Lin et al., 2013).

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 (Attfield et al., 2011). We argue here that student engagement indeed may be quantified on a second-by-second basis by measuring electroencephalography (EEG) responses during exposure to media stimuli.

Traditional approaches to measuring engagement (O'Brien and Toms, 2013) are based on capturing user behaviour via user interfaces, self-report, or manual annotation. However, increasingly the tools of cognitive neuroscience are employed (Szafir and Mutlu, 2013). Here we use neural responses to video to quantify students' engagement with the content. The work is based on the approach developed by Dmochowski et al. (2012). 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 (Ki et al., 2016).

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. To relate the finding on engagement measured with neural activity (Dmochowski et al., 2012, 2014; Ki et al., 2016) to real-time classroom engagement several issues must be addressed, including: Is it possible to reproduce the detection of engagement under the adverse conditions of a classroom? Is the detection robust to inter-student variability of the spatial information processing networks? Can engagement detection be performed with equipment that is both comfortable and low enough costs to make it a realistic technology for schools? We have designed a set of experiments to address these issues. Hence, a main goal of the present work is to determine if student engagement can be quantified in a real-time manner using recordings of brain activity in a classroom setting using a low-cost, portable EEG system – the Smartphone Brain Scanner (Stopczynski et al., 2014b). On the robustness of the detection scheme we report on both theoretical and experimental investigations. First we show mathematically that the detection algorithm is surprisingly robust to the spatial structure of the students brain networks. Secondly we relate the engagement signal to a very basic visual response, in particular, we demonstrate that the strength of visually evoked activity is modulated by narrative coherence of the video stimulus – reflecting the engagement of the group of students.

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 such that the narrative was lost ("Scrambled"). This condition aims 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 ($N = 2 * 9$) watched the original videos on a screen in classroom (Figure 1a, "Joint 1" and "Joint 2"), with sound projected through loudspeakers. An attempt was made to create similar viewing conditions as in the individual viewing experiment, i.e., lights were dampened and the projected image produced approximately the same field-of-view (see supplementary materials) for the joint viewing subject. The central question was whether the viewing condition (i.e., in a group versus individually) influences the level of neural response reliability across subjects.

Stimuli: The first video clip is a suspenseful excerpt from the short film *Bang! You're Dead* (1961) directed by Alfred Hitchcock. It was selected here because it is know to elicit highly reliable brain activity across subjects in fMRI (Hasson et al., 2004) as well as EEG (Dmochowski et al., 2012). Our second stimulus was a clip from *Sophie's Choice*, earlier used to study fMRI activity in the context of emotionally salient naturalistic stimuli (Raz et al., 2012). A third non-narrative control video was recorded in a Danish metro station with several people transported quietly on an escalator. Each video clip had a length of approximately 6 minutes and was shown twice to each subject. For each viewing the order was randomized, but the same 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* following previous research (Hasson et al., 2008; Dmochowski et al., 2012)).

Subjects: A total of 42 female subjects were recruited for this study (mean age: 22.4y, age range: 18-32y), who gave written informed consent prior to the experiment. Among the 42 recordings, 8 subjects were excluded due to unstable wireless communication that precluded proper synchronization of the data across subjects (9 in individual viewing, 3 in joint viewing). To enable fair statistical comparisons, we equalized the number of participants between groups by selecting 7 recordings, which was the number of fully synchronized

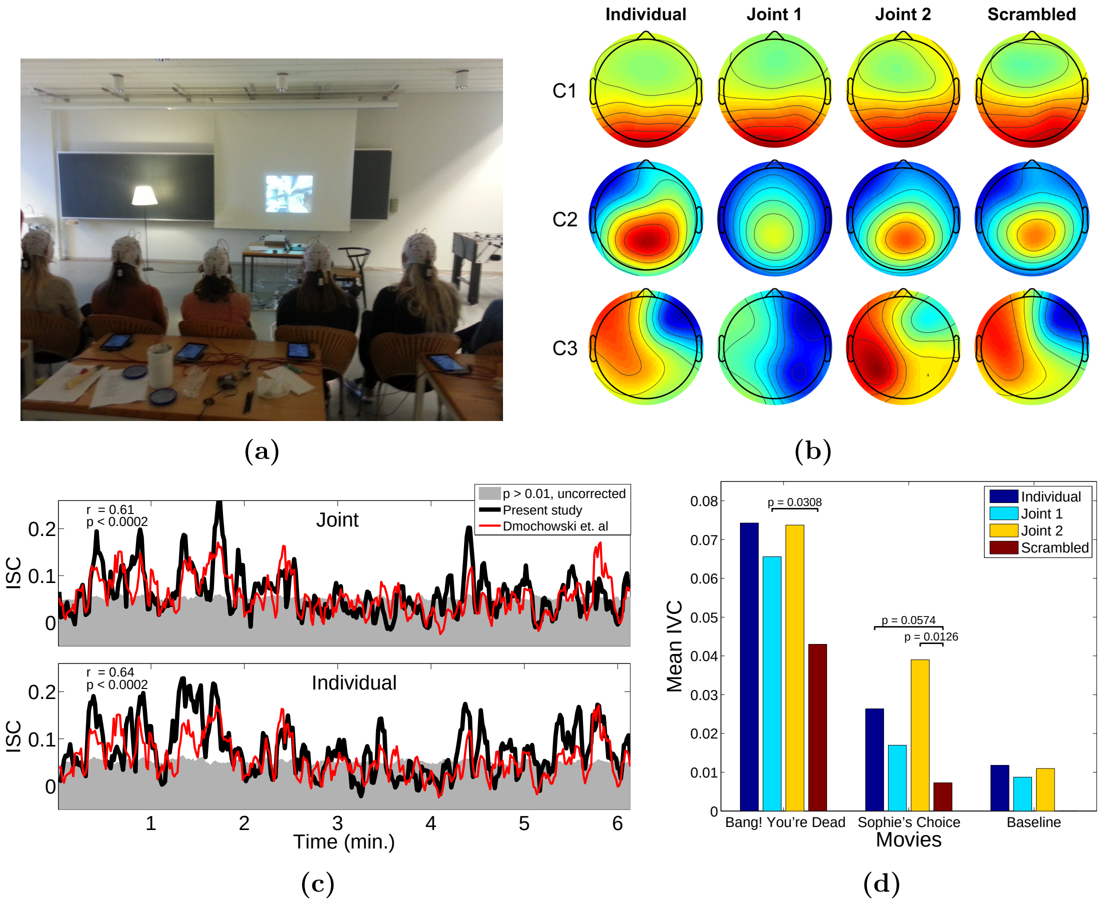

Figure 1: Joint recording of neural responses to a naturalistic stimulus in a classroom setting. **(a)** Experimental setup for joint viewings. Subjects seen from the back before viewing the films. All subjects were placed on a line to induce a cinema-like experiences. **(b)** The scalp projections of the first three components obtained from the correlated component analysis of each of the four subject groups, watching *Bang! You're Dead* the first time. 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*), jointly in a classroom as in (a) (*Joint 1* and *Joint 2*). For each pattern, the polarity was normalized so the value at Cz is positive. **(c)** Comparison between the ISC obtained by Dmochowski et al. (2012) and the present study. The ISCs are calculated from the first viewing of *Bang! You're Dead*. The grey area signifies correlation corresponding to $p > 0.01$ in the permutation statistics. **(d)** Mean IVC calculated for the strongest correlated component averaged in time. A block permutation test (block size $B = 25\ s$) has been employed to estimate statistical significant differences between viewing conditions. Note that the *Scrambled* group did not watch the baseline movie.

recordings present in all groups.

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 downsampled to 1 Hz by selecting the maximum ALD for each 1 s interval (see figure S1 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 of 2.5 s^2. This down sampling and smoothing intends to match the temporal resolution of the ALD to that of the time-resolved ISC computation (5 s sliding window with 1 s intervals).

Portable EEG – Smartphone brain scanner: Research grade EEG equipment is costly, time-consuming to set up, and immobile. However, recently consumer grade EEG equipment has appeared at reduced price and increased comfort. Here we use the modified 14 channel system, 'Emocap', based on the EEG Emotiv EPOC headset. For details and validation, see (Stopczynski et al., 2014a,b). The present version 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. The resulting timing uncertainty was measured to be less than 16 ms. The recorded EEG was bandpass filtered 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 (Delorme and Makeig, 2004; Viola et al., 2009)).

Correlated component analysis to measure ISC and IVC: Correlated components analysis (CorrCA, described below) was used 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). To investigate the temporal development of ISCs and IVCs, they were calculated in 5 s windows with 80 % overlap, resulting in a 1-second resolution. In order to evaluate the statistical relevance of the correlations, we employed a simple permutation test ($P = 5000$ permutations)(Dmochowski et al., 2012). When testing for differences in average IVC between conditions, a block permutation test (block size $B = 25$ s, $P = 5000$ permutations) was used in order to account for temporal dependencies.

CorrCA was developed by Dmochowski et al. (2012), as a constrained version of Canonical Correlation Analysis (CCA). Given two multivariate spatio-temporal time series, $\{\mathbf{X}_1, \mathbf{X}_2\} \in \mathbb{R}^{D \times N}$, 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, $\mathbf{y}_1 = \mathbf{X}_1^\intercal \mathbf{w}_1$ and $\mathbf{y}_2 = \mathbf{X}_2^\intercal \mathbf{w}_2$. The weights are calculated using two eigenvalue equations, with the constraint that the components within each view are uncorrelated (Hardoon et al., 2004). 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;

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

where, $\mathbf{R}_{ij} = \frac{1}{N}\mathbf{X}_i \mathbf{X}_j^\intercal$, is the sample covariance matrix (Dmochowski et al., 2012). To illustrate the spatial distribution of the underlying physiological activity, the forward models, $\mathbf{A} = \mathbf{RW}(\mathbf{W}^\intercal \mathbf{RW})^{-1}$, are used (Parra et al., 2005).

Robustness to inter-subject variations in the spatial brain structure is a basic question when applying CorrCA to classroom data. In particular CorrCA assuming that subjects' spatial networks are identical could be challenged by inter-individual differences, however, it turns out to be surprisingly robust to such variability. To demonstrate this, we briefly analyse a 'worst case' scenario in which the true mixing weights form *orthogonal* vectors. The observations are assumed to consist of a single true signal, $\mathbf{z}$, mixed into D dimensions with additive Gaussian noise; $\mathbf{X}_1 = \mathbf{a}_1 \mathbf{z}^\intercal + \boldsymbol{\epsilon}$, $\mathbf{X}_2 = \mathbf{a}_2 \mathbf{z}^\intercal + \boldsymbol{\epsilon}$. Given a large sample, the covariance matrices are given as $\mathbf{R}_{11} = P \cdot \mathbf{a}_1 \mathbf{a}_1^\intercal + \sigma^2 \mathbf{I}$, $\mathbf{R}_{12} = P \cdot \mathbf{a}_1 \mathbf{a}_2^\intercal$, 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. (1) can then be written as

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

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

$$(\mathbf{R}_{11} + \mathbf{R}_{22})^{-1} (\mathbf{R}_{12} + \mathbf{R}_{21}) = \frac{P}{2\sigma^2 + P} (\mathbf{a}_1 \mathbf{a}_2^\intercal + \mathbf{a}_2 \mathbf{a}_1^\intercal) \tag{3}$$

An eigenvector of matrix (3) 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. Thus, even if the weights of the two views are orthogonal, CorrCA still identifies the relevant time series using a weighted sum of the two true weights. Given the joint activation times series, subject-specific forward models (scalp maps) can be inferred by linear regression of the individual observations against the common activation time series, so-called inverse correlation.

Results

To monitor attentional engagement we used video stimuli as they provide a balance between realism and reproducibility (Hasson et al., 2004). We recorded EEG activity using the Smartphone brain

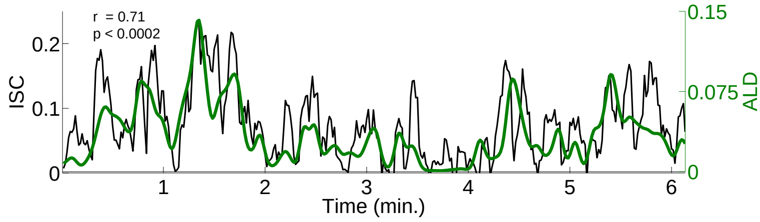

Figure 2: The ISC of Correlated Component 1 is temporally correlated with luminance differences of the stimulus, consistent with the view that ISC is driven by attentional modulation of visual evoked responses. Data computed from the neural responses to *Bang You're Dead*.

scanner while subjects watched short video clips of approximately 6 minute duration, either individually or in a group setting (Figure 1a). Three video clips were presented in random order and repeated twice. To measure reliability of EEG responses, we then computed inter-subject correlation (ISC) and inter-viewing correlation (IVC) using established methods (see Methods). For both the joint and individual recording scenarios the fluctuations of ISC closely reproduce results obtained previously in a laboratory setting for one of the clips (Figure 1c and Table 1). This clip is a suspenseful excerpt from a short film directed by Alfred Hitchcock (*Bang! You're Dead*, 1961). It was selected because it is known to effectively synchronize brain responses across viewers (Hasson et al., 2008; Dmochowski et al., 2012).

	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 results obtained in a laboratory setting (Dmochowski et al., 2012) 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 first and second viewings (v1,v2) and the inter-viewing correlation (IVC) measures similarity within-subject between the two views. Values are calculated using the first correlated component recorded while watching *Bang! You're dead*. **: $p < 0.01$, *: $p < 0.05$.

To determine if the portable equipment, which uses only 14 electrodes, can detect varying levels of engagement, a second group of subjects watched the same video clips individually, but now with scenes scrambled in time. This manipulation aims to emulate reduced viewer engagement and is known to reduce reliability of neural responses (low IVC) while maintaining identical low-level stimulus features (Dmochowski et al., 2012). Despite using consumer-grade EEG we find that IVC is significantly above chance for a large fraction of the original engaging clip, but this drops dramatically when the scenes are scrambled in time (mean IVC, Figure 1d, $p < 0.01$, for *Bang! You're Dead*). Also, as expected, the baseline video shows no or little synchronisation, which is further evidence that the IVC is affected by viewer engagement.

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

A further test of the stability of the technique is provided by the spatial patterns of the neural activity that drives these reproducible responses. The technique we used to calculate ISC and IVC measures correlation, not of individual EEG electrodes, but of components of the EEG, i.e. linear combinations of electrodes as used in other component extraction techniques, such as independent

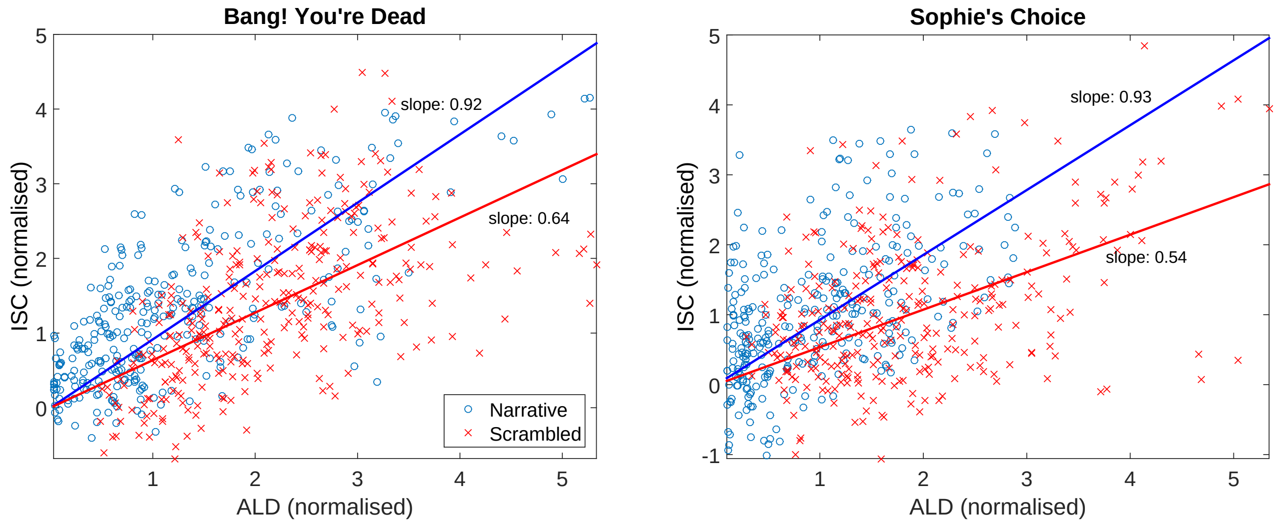

Figure 3: Dependence of ISC on ALD for different engagement conditions. Each point indicates instantaneous ISC and ALD calculated from the visual stimulus. It is evident that moments with higher luminance fluctuations (hight ALD) result in higher correlation of brain activity across subjects (high ISC). Slope indicates the strength of ISC for a given ALD value. 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 video (red).

components or common spatial patterns (Parra and Sajda, 2003; Koles et al., 1990). The strongest two correlated components show a stable pattern of activity across the different groups and recording scenarios (Figure 1b).

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 an film audience follow a specific pattern after a scene change, activating the dorsal pathway (Unema et al., 2005). Likewise, it is also known that stimulus in the form of flashing images elicits visual evoked responses (VEP), which are modulated in amplitude by the luminance (Armington, 1968). When recorded with EEG, the spatial distribution of the early VEP at 100ms (P100) is similar to the scalp maps of the first correlated component (C1 in figure 1b) (Johannes et al., 1995; Sandmann et al., 2012). To estimate how much of the response is driven by such a basic mechanism, 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 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).

	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 2: Correlation coefficients between the ALD and the ISC for the two viewings (v1,v2) as well as the IVC for the first correlated component C1. The correlation is presented for *Bang You're dead* and *Sophie's Choice* for the *Individual* and *Scrambled (Scr)* groups. **: $p < 0.01$, *: $p < 0.05$.

Figure 2 and Table 2 show that there is a significant correlation between the ISC and the ALD for both *Bang! You're Dead* and *Sophie's Choice*. This suggests that this portion of the correlated activity (C1) may indeed be driven by low-level visual evoked responses. However, the level of attention modulates the *amplitude* of the ISC time course, in that the ISC elicited by the scrambled stimulus elicits a much lower (but still presumably driven by attention) ISC. Modulation of VEPs by attention has been previously documented (Johannes et al., 1995). We quantify the dependence on attention by comparing the sensitivity (slope) of ISC to ALD in both the normal and scrambled conditions (Figure 3). For both movies we found significant reductions of the ISC/ALD slope in the scrambled version ($p < 0.01$; block permutation test, with block size $B = 25\ s$).

Fig. 2 indicates that the interval in *Bang! You're dead* with the highest and most sustained ISC (around 1:20 to 1:50) coincides with the interval of the clip in which scene changes are most numerous. Note that fast-paced cutting is a known cinematographic tool used by Hitchcock to induce suspense and thereby increase the attention of the viewer (Bordwell, 2002).

Conclusion

We have demonstrated that student engagement in media stimuli may be measured using EEG in a classroom setting. For educational technology cost and robustness are a 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. Mathematically, we showed that the our detection scheme, CorrCA, is robust to inter-subject variability in spatial configurations of brain networks. We provided evidence that salient aspects of the attentional modulation earlier 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 and predicted viewer engagement. We also found that luminance fluctuations drive a significant portion of this correlation, and that the strong dependence of inter-subject correlation on narrative coherence of the stimulus is at least partly due attentional modulation of visual evoked responses. The evidence presented that such a basic mechanism is at play further add to the robustness of the approach.

References

Armington, J. C. (1968). The electroretinogram, the visual evoked potential, and the area-luminance relation. *Vision research*, 8:263–276.

Attfield, S., Kazai, G., Lalmas, M., and Piwowarski, B. (2011). Towards a science of user engagement. In *WSDM Workshop on User Modelling for Web Applications*. ACM International Conference on Web Search And Data Mining.

Bordwell, D. (2002). Intensified continuity: visual style in contemporary American film. *Film Quarterly*, 55(3):16–28.

Delorme, A. and Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21.

Dmochowski, J. P., Bezdek, M. a., Abelson, B. P., Johnson, J. S., Schumacher, E. H., and Parra, L. C. (2014). Audience preferences are predicted by temporal reliability of neural processing. *Nature Communications*, 5:1–9.

Dmochowski, J. P., Sajda, P., Dias, J., and Parra, L. C. (2012). Correlated components of ongoing EEG point to emotionally laden attention - a possible marker of engagement? *Frontiers in human neuroscience*, 6(May):112.

Hardoon, D. R., Szedmak, S., and Shawe-taylor, J. (2004). Canonical correlation analysis; An overview with application to learning methods. *Neural computation*, 16(12):2639–2664.

Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., and Heeger, D. J. (2008). Neurocinematics: The neuroscience of film. *Projections*, 2(1):1–26.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., and Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *science*, 303(5664):1634–1640.

Johannes, S., Münte, T. F., Heinze, H. J., and Mangun, G. R. (1995). Luminance and spatial attention effects on early visual processing. *Cognitive Brain Research*, 2:189–205.

Ki, J. J., Kelly, S. P., and Parra, L. C. (2016). Attention strongly modulates reliability of neural responses to naturalistic narrative stimuli. *The Journal of Neuroscience*, 36(10):3092–3101.

Koles, Z. J., Lazar, M. S., and Zhou, S. Z. (1990). Spatial patterns underlying population differences in the background eeg. *Brain topography*, 2(4):275–284.

Lahnakoski, J. M., Glerean, E., Jääskeläinen, I. P., Hyönä, J., Hari, R., Sams, M., and Nummenmaa, L. (2014). Synchronous brain activity across individuals underlies shared psychological perspectives. *NeuroImage*, 100:316–24.

Lin, C.-T., Huang, K.-C., Chuang, C.-H., Ko, L.-W., and Jung, T.-P. (2013). Can arousing feedback rectify lapses in driving? prediction from eeg power spectra. *Journal of neural engineering*, 10(5):056024.

O'Brien, H. L. and Toms, E. G. (2013). Examining the generalizability of the user engagement scale (ues) in exploratory search. *Information Processing & Management*, 49(5):1092–1107.

Parra, L. and Sajda, P. (2003). Blind source separation via generalized eigenvalue decomposition. *The Journal of Machine Learning Research*, 4:1261–1269.

Parra, L. C., Spence, C. D., Gerson, A. D., and Sajda, P. (2005). Recipes for the linear analysis of EEG. *NeuroImage*, 28(2):326–41.

Raz, G., Winetraub, Y., Jacob, Y., Kinreich, S., Maron-Katz, A., Shaham, G., Podlipsky, I., Gilam, G., Soreq, E., and Hendler, T. (2012). Portraying emotions at their unfolding: a multilayered approach for probing dynamics of neural networks. *NeuroImage*, 60(2):1448–61.

Ringach, D. L., Hawken, M. J., and Shapley, R. (2002). Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences. *Journal of vision*, 2(1):2–2.

Sandmann, P., Dillier, N., Eichele, T., Meyer, M., Kegel, A., Pascual-Marqui, R. D., Marcar, V. L., Jäncke, L., and Debener, S. (2012). Visual activation of auditory cortex reflects maladaptive plasticity in cochlear implant users. *Brain*, 135:555–568.

Stopczynski, A., Stahlhut, C., Larsen, J. E., Petersen, M. K., and Hansen, L. K. (2014a). The smartphone brain scanner: a portable real-time neuroimaging system. *PloS one*, 9(2):e86733.

Stopczynski, A., Stahlhut, C., Larsen, J. E., Petersen, M. K., and Hansen, L. K. (2014b). The smartphone brain scanner: a portable real-time neuroimaging system. *PloS one*, 9(2):e86733.

Szafir, D. and Mutlu, B. (2013). Artful: adaptive review technology for flipped learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1001–1010. ACM.

Unema, P. J., Pannasch, S., Joos, M., and Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition*, 12(3):473–494.

Viola, F. C., Thorne, J., Edmonds, B., Schneider, T., Eichele, T., and Debener, S. (2009). Semi-automatic identification of independent components representing EEG artifact. *Clinical Neurophysiology*, 120(5):868–877.

Acknowledgements

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 statement

ATP, SK, JD, LP and LKH designed research; SK, ATP and LKH performed research; ATP, SK, JD, LP, and LKH contributed analytical tools; ATP, SK, LKH analysed data; ATP, SK, JD, LP, and LKH wrote the paper.

Additional information

Competing financial interests The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.