Tracking individual auditory attention during learning via EEG-based neural envelope tracking

Davina Van den Broek[1], Elien Bellon[1], Tianyi Li[2], Simon Geirnaert[2], Alexander Bertrand[2] and Bert De Smedt[1]

1 Faculty of Psychology and Educational Sciences, Parenting and Special Education Research Group, Leuven, Belgium

2 Department of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, KU Leuven, Belgium

Corresponding author: Davina Van den Broek – davina.vandenbroek@kuleuven.be

Funding information: *This study was funded by an interdisciplinary research grant (ID-N) of the KU Leuven Research Fund (IDN/23/006). E.B. (12C9523N) and S.G. (1242524N) were supported by a postdoctoral fellowship from the Flemish Research Foundation (FWO). The funders played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript.*

Abstract

Attention is essential for classroom learning, yet measuring it during learning remains challenging. Traditional behavioural measures are often subjective and lack the precision needed to capture moment-to-moment shifts between on-topic and off-topic attention. This study explores the potential of EEG (Electroencephalography)-based neural envelope tracking (NET) as an indicator of external on-topic auditory attention to the teacher's voice. Unlike more traditional EEG-measures of attention, NET considers where attention is directed to by correlating the teacher's voice with the EEG-decoded stimulus. EEG data were collected while participants (*n* = 30) watched a video lecture on neuromyths that included attention manipulations (e.g., distraction, enhancement). Participants also provided self-reported attention ratings and completed a performance test on its content. The results indicated that NET was sensitive to decreases in external on-topic attention during distractions, supporting the potential of NET as an indicator of external on-topic auditory attention during learning.

1. Introduction

The concept of attention has been the focus of extensive debate, reflecting its complexity and the absence of a universally accepted definition. Nevertheless, there is a broad consensus that attention should be regarded as an umbrella term for multiple different but related processes (Anderson, 2011; Cohen, 2014; Di Lollo, 2018; Hommel et al., 2019; Styles, 2006). To better understand attention in educational contexts, Keller and colleagues (2020) proposed a framework specifically tailored to classroom learning. It distinguishes two dimensions: (1) internal versus external attention and (2) on-topic versus off-topic attention. This results in four distinct attentional states: (1) internal on-topic attention, (2) external on-topic attention, (3) internal off-topic attention and (4) external off-topic attention. This framework enables characterisation of students' attentional states at any given moment during classroom learning, which is important because variations in on-topic and off-topic attention are associated with differences in learning outcomes (Olney et al., 2015; Keller et al., 2020). Capturing these momentary shifts, however, remains challenging. Existing methods range from behavioural to physiological measures, each offering valuable but incomplete insights into specific aspects of attentional processes (Booth et al., 2023). In this study, we focus specifically on the measurement of external on-topic auditory attention directed toward the educational content delivered by the teacher in a video lecture.

Common behavioural measures of external on-topic attention include self-reports, observations, and performance tests. Self-reports are typically collected through questionnaires, and a common approach involves administering those self-reports at the end of, for example, a lecture, allowing participants to reflect on their overall attentiveness (Blasiman et al., 2018; Hobbis & Lavie, 2024). However, because they require participants to assess their external on-topic attention retrospectively over an extended period, they fail to capture moment-to-moment fluctuations. To address this, some studies used a probe-caught method, in which participants are asked during a lecture (Bühler et al., 2024; Forrin et al., 2021; Risko et al., 2013; Smallwood et al., 2007; Varao-Sousa & Kingstone, 2019) or while reading a text (Smallwood et al., 2007) what they were thinking about just before the probe questions appeared. Response options vary widely, ranging from two categories (on-task vs. off-task) to seven categories (e.g., on-task, day-dreaming, personal concerns), and often include an indication of whether the participant was aware or unaware of the attention lapse. The frequency of probes also differs substantially across studies, partly depending on the length of the lecture or text, but even within similar contexts there is a considerable variability. However, this method only provides information about attention immediately prior to the probe question, and may disrupt periods of high attention, which is to be avoided when measuring attention (Keller et al., 2020). A potential solution is the use of another self-report method, namely, the self-caught method using clickers, in which participants indicate attention lapses as they occur (Bunce et al., 2010; Hlas et al., 2019). In both studies, participants received a clicker with three buttons, each representing a different duration of distraction, such as 1 minute or less, 2-3 minutes, or 5 minutes or more, allowing participants to report how long they had been distracted from the course content. Although these self-report methods offer valuable insights into external on-topic attention during learning, they rely on participants' self-awareness and their interpretation of external on-topic attention, creating a risk of biased self-reporting.

Another common behavioural method for assessing external on-topic attention is observation (Grammer et al., 2021; Keller et al., 2020; Rapport et al., 2009). In these studies, attention is operationalised through observable indicators of attentive behaviours, such as gaze and body positioning. Observers code behaviour either in predefined intervals as attentive or inattentive (Rapport et al., 2009) or based on the duration of specific inattentive cues (Grammer et al., 2021). However, because observations rely on the interpretation of behaviour and gaze by the observer, they are just as subjective as self-reports (Bradbury, 2016; Keller et al., 2020).

While self-reports and observations provide direct measures of external on-topic attention, performance tests are often used as indirect indicators of external on-topic attention (Bevilacqua et al., 2019; Boudewyn & Carter, 2018; Davidesco et al., 2023; Dikker et al., 2020; Zeamer & Fox Tree, 2013). In all these studies, participants completed a test with items about the content of the lecture they just attended, and higher scores are assumed to reflect higher levels of attention. However, performance depends on many factors beyond attention, such as prior knowledge and motivation, meaning it cannot be seen as a pure measure of external on-topic attention alone.

The major limitations of all the abovementioned behavioural measures are their reliance on interpretation and/or self-awareness, their indirect nature, and their inability to capture the rapid fluctuations of attention during learning. Neurophysiological measures have been suggested as promising, more objective tools for assessing attention, as they allow for the monitoring of cognitive processes during learning rather than afterwards, without interrupting attention by asking questions or relying on interpretation and self-reports (De Smedt, 2018; Mayer, 2017). Especially electroencephalography (EEG), a non-invasive brain imaging technique with high temporal resolution, has been suggested to overcome the limitations of behavioural methods as it has the possibility to capture momentary fluctuations of attention in the classroom without being affected by subjective interpretation (Malik & Amin, 2017; Mayer, 2017).

A first commonly used EEG measure of attention in the classroom is brain-to-brain synchrony, observed either between students or between students and their teacher. This phenomenon occurs because individuals exhibit similar neural activity during shared learning experiences, and this synchrony increases with greater attention (Bevilacqua et al., 2019; Cohen et al., 2018; Davidesco, 2020; Davidesco et al., 2023; Dikker et al., 2017; Dmochowski et al., 2012; Ki et al., 2016; Liu et al., 2025; Madsen & Parra, 2022; Poulsen et al., 2017). Davidesco et al. (2023) examined both student-student and student-teacher brain-to-brain synchrony. In their experiment, a group of four students attended four mini-lectures (~ 7 minutes). They found that increased brain-to-brain synchrony between student dyads predicted better immediate and delayed learning outcomes, while student-teacher synchrony only predicted immediate learning outcomes. Bevilacqua et al. (2019) similarly investigated student-student and student-teacher brain-to-brain synchrony in a real-world high school biology classroom. Unlike Davidesco et al. (2023) they could not find a direct link between brain-to-brain synchrony and immediate learning outcomes. Although brain-to-brain synchrony reflects shared neural processing, it is not exclusively driven by attention, it is also affected by other factors such as the nature of the stimuli and social dynamics (Bevilacqua et al., 2019; Davidesco, 2020; Dikker et al., 2017; Liu et al., 2025). Two more key challenges are faced when using this measure for external on-topic attention. First, because brain-to-brain synchrony reflects a shared experience, synchrony may also increase when students are jointly distracted by the same external stimulus. Second, this method requires EEG data from multiple individuals, therefore it cannot be used to track external on-topic attention at an individual level.

To capture attention at an individual level with EEG, alpha power has been frequently used (Boudewyn & Carter, 2018; Foxe & Snyder, 2011; Levy, Shadi, et al., 2025; Matsuo et al., 2024; O’Connell et al., 2009), with lower alpha power reflecting increased external attention (Boudewyn & Carter, 2018; Compton et al. 2019; Grammer et al., 2021; Ki et al., 2016; Polychroni et al., 2022). Grammer et al. (2021) used alpha power to compare attention across instructional contexts. In this study, a small group of students participated in four different instructional activities: whole-group lecture, video watching, group discussion, and independent work. Alpha power was significantly higher, indicating lower attention, in teacher-led activities compared to student-initiated activities. Similarly, Dikker et al. (2020) examined differences between educational videos and traditional lectures using alpha power, self-reports of attention and learning outcomes. Data were collected from 22 high school students across multiple biology classes. The findings revealed that alpha power was lowest during educational videos compared to lectures, aligning with higher self-reports and better test performance for educational videos. Ki et al.

(2016) examined attention using alpha power across audiovisual versus auditory-only stimuli, each presented once with instructions to attend and once with the instructions to count down from 1000 in steps of 7. The results indicated higher alpha power, indicating lower attention, in the counting condition and the auditory-only condition, as anticipated. However, while alpha power can indicate whether a student is generally attentive, it cannot reveal the focus of attention. For instance, a student might exhibit low alpha power during a lecture, suggesting high attention, but the student's attention could be directed towards something other than the teacher's speech, such as their smartphone. Consequently, despite the low alpha power, the student may miss what the teacher is saying, leading to a misinterpretation of their attentiveness to the lecture. Addressing this key limitation is essential for understanding how external on-topic attention contributes to learning outcomes.

To address the key challenges of brain-to-brain synchrony and alpha power, neural envelope tracking (NET) provides a stimulus-specific and individual measure of external on-topic auditory attention, allowing to determine whether students are attending to the teacher's speech. NET is based on the assumption that when people actively listen to a specific person, their brain signals synchronise with the short-term amplitude fluctuations of the speech signal (i.e., the envelope of the speech signal), and this coupling is assumed to decrease when not attending to the speech. To quantify external on-topic auditory attention using NET, a correlation is calculated between the actual speech envelope and a reconstructed speech envelope from the brain responses using a pre-trained EEG decoder. In previous studies using NET, the correlation between these two tends to reflect the individual's auditory attention to the stimulus (Mirkovic et al., 2015; O'Sullivan et al., 2015; Roebben et al., 2024; Straetmans et al., 2024; Vanthornhout et al., 2019).

This method has been employed by multiple researchers for selective attention decoding in competing speech paradigms, in which participants attend to one speech stream and ignore the other. In the studies by O'Sullivan et al. (2015) and Mirkovic et al. (2015) participants were asked to listen to two different stories simultaneously and to focus on one of the stories. In both studies, NET was able to reliably identify which story participants were attending to. Straetmans et al. (2024) conducted a similar study but in more complex environments by adding background noise to the paradigm and by conducting the research in an everyday life setting, such as sitting in a hallway or walking on the street. Even in these more complex environments, it was possible to identify the attended speaker among other speakers using NET. Other studies have used NET for absolute attention decoding, examining fluctuations in auditory attention to a single stimulus. The studies by Roebben et al. (2024) and Vanthornhout et al. (2019) included a condition designed to induce attentional disengagement, in which participants were instructed to ignore the auditory stimulus while performing a different task (e.g., watching a silent movie, reading a text) and compared it with an active listening condition in which full attention is paid to the auditory stimulus. They found that NET was higher during active listening, indicating its sensitivity to attentional fluctuations within individuals.

To the best of our knowledge, only three studies have explored NET for absolute attention decoding in an educational context (Levy, Hackmon, et al., 2025; Levy, Korisky, et al., 2025; Levy, Shadi, et al., 2025). In the first study (Levy, Korisky, et al. 2025), the impact of background noise on performance and NET was investigated. Participants viewed short (~40s) mini-lectures in a VR classroom while EEG was recorded. The mini-lectures were divided in three conditions: quiet, continuous background noise, and intermittent background noise. After each lecture, they answered four multiple-choice questions. Performance was slightly, but not significantly, lower in the noise condition compared to the quiet condition, and significantly worse in the intermittent noise condition compared to the continuous noise condition. Similarly, NET was significantly reduced in the noise conditions compared to the quiet condition, with the most pronounced reduction in the intermittent noise condition. The second study (Levy, Hackmon, et al. 2025), used a similar design with 30 mini-lectures, 22 of which included sporadic background noise (e.g., human and artificial sounds). Participants again answered four multiple-choice questions after each lecture. The study compared attention and distractibility in individuals with and without AD(H)D. No significant performance differences were found between quiet and noise conditions, or between groups. NET was slightly lower for the AD(H)D group in the noise condition, suggesting reduced external on-topic auditory attention in the presence of background noise. In the last study (Levy, Shadi, et al., 2025), a 35-minute lecture was used to investigate the impact of situational interest and background noise on NET, among other outcome measures. The lecture was divided into 63 segments, lasting 23–40 seconds each. In between these segments, participants rated their situational interest on a 7-point Likert scale and after every three segments, participants answered three multiple-choice questions on the content of the lecture. Each segment was assigned to one of the conditions: silent, continuous background noise or intermittent background noise. The results indicated that both background noise and situational interest significantly influenced NET, however, the impact of situational interest was stronger.

Evidence for the applicability of NET to assess external on-topic auditory attention in an educational context remains limited. To properly evaluate its potential in classroom contexts, studies should incorporate educationally relevant stimuli to obtain results that are truly reflective of students' attentional processes during classroom learning (van Atteveldt et al., 2018). Most prior research has relied on non-educational content, such as children’s stories (Mirkovic et al., 2015; O’Sullivan et al., 2015; Roebben et al., 2024; Straetmans et al., 2024; Vanthornhout et al., 2019), paradigms with competing speakers (O’Sullivan et al., 2015; Mirkovic et al., 2015; Straetmans et al., 2024), or stimuli shorter than a typical classroom lecture (Levy, Hackmon, et al., 2025; Levy, Korisky, et al., 2025). These approaches do not reflect the complexity of classroom learning, where students must sustain attention and engage with educational material over extended periods.

Against this background, we examined whether NET could be used as an indicator of external on-topic auditory attention to the teacher’s voice. Participants watched a 65-minute video lecture while EEG was recorded. During this lecture, we included multiple interventions to manipulate the external on-topic attention of the participants (Figure 1). Specifically, we tried to enhance attention by displaying a large red exclamation mark and a red frame around the video, with a prior instruction that this indicates an important part of the lecture. In order to create distraction, a funny animal video or disturbing background noise, such as in-class background noise or construction works, was incorporated. To induce attentional disengagement, participants were instructed to ignore the video lecture and perform an alternative task. Based on prior studies, we hypothesised that NET would be reduced during conditions designed to disrupt (distracting condition) or withdraw (disengagement condition) attention from the video lecture (Levy, Korisky et al., 2025; Levy, Shadi, et al., 2025; Roebben et al., 2024; Vanthornhout et al., 2019). In addition to EEG, we collected self-reported attention at fixed time points using five 10-point Likert-scale questions and assessed learning with a post-lecture performance test. This allowed us to validate NET by examining its associations with established behavioural methods, which we hypothesised would reflect external on-topic attention fluctuations induced by our manipulations (Blasiman et al., 2018; Levy, Korisky et al., 2025; Zeamer & Fox Tree, 2013). Finally, we compared NET to more general EEG-based indicators of attention in the classroom, namely brain-to-brain synchrony and alpha power. We hypothesised that NET would be more sensitive to fluctuations in students’ attention to the teacher as compared to these measures. Specifically, we expected lower NET during the distracting condition, while brain-to-brain synchrony would likely remain high because participants experienced the same distractions simultaneously, maintaining shared neural responses. Similarly, we expected lower NET in the attentional disengagement condition but did not expect alpha power to clearly reflect this drop in attention to the teacher’s voice, since participants remain engaged in an external task. Although a shift in alpha power is expected, its direction is unclear since both listening and the other task (e.g., reading and math exercises) require attentional engagement (Boudewyn & Carter, 2018; Compton et al. 2019; Grammer et al., 2021; Ki et al., 2016; Polychroni et al., 2022).

2. Method

2.1. Participants

Data were collected from a sample of 30 participants (7 males), aged between 18 and 35 (M_{age} = 24 years; *SD* = 3 years). All participants were native Dutch speakers with normal to corrected-to-normal vision and no hearing difficulties. None of the participants reported a deficit in attention. The participants' backgrounds varied considerably in terms of their studies and careers, although nearly all were either pursuing or had completed a higher education degree.

Participants received monetary compensation for their participation. Written consent was obtained from all participants, and the procedure was approved by the Social and Societal Ethics Committee (SMEC) of KU Leuven (G-2023-7126).

2.2. Design

The data collection consisted of two parts (Figure 1). First, EEG data were collected while participants watched an online lecture. During breaks in the lecture, participants answered five questions about their external on-topic attention. Second, participants were asked to complete a performance test with 37 items about the video's content.

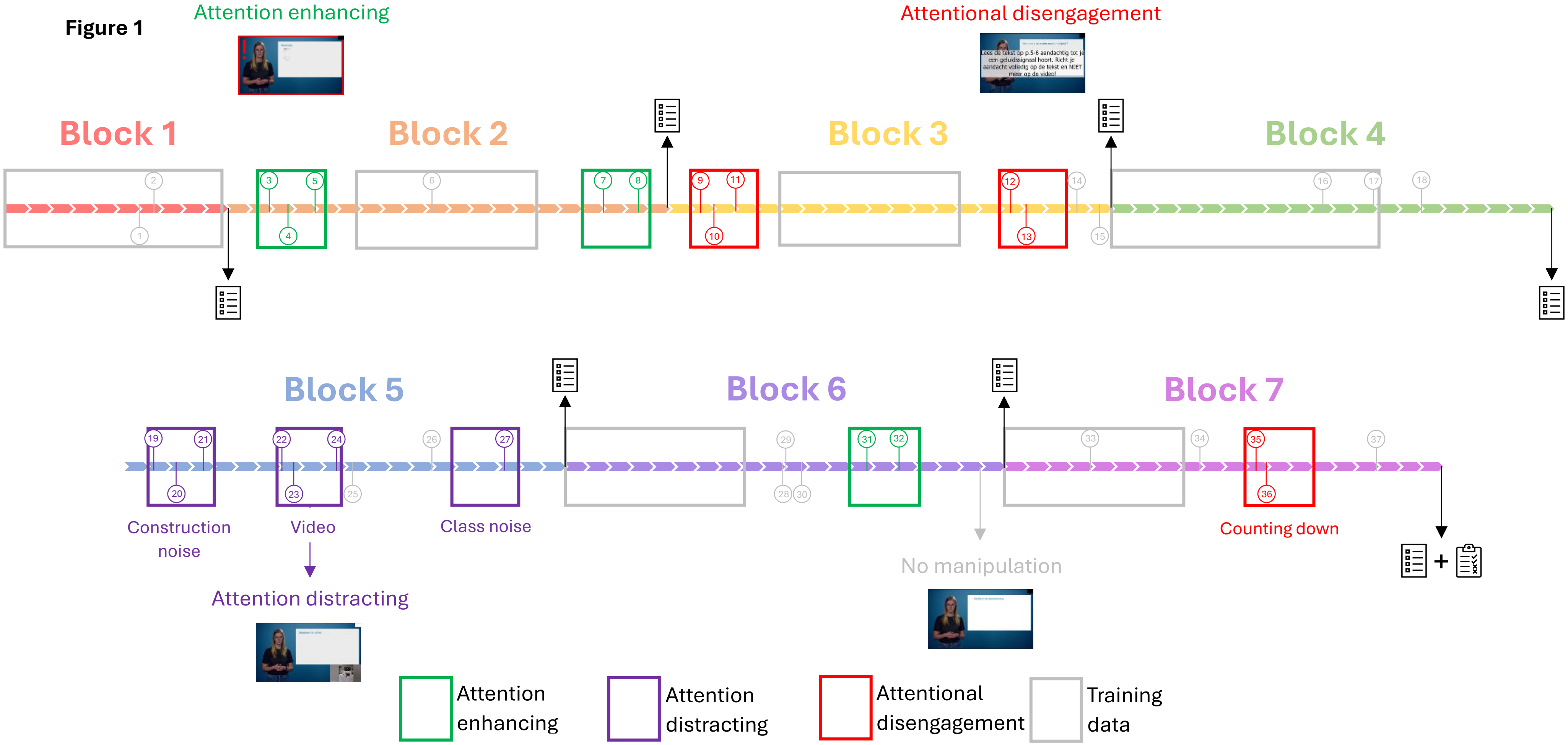
The online lecture covered the topic of neuromyths (Howard-Jones, 2010; Jarrett, 2015) and had a total duration of 65 minutes. The video depicted a teacher and a PowerPoint presentation containing only images (Figure 1). As the aim was to measure external on-topic auditory attention, all information was given orally, with the images only supporting the spoken content. These images were not intelligible without the accompanying speech. A schematic representation of the video is provided in Figure 1. The video was divided into seven blocks: an introduction (5 minutes) and the six subsequent blocks, each addressing a different neuromyth (10 minutes each). The following neuromyths were included: (1) study pills help you study better and become smarter, (2) we only use 10% of our brains, (3) there are critical periods in child development; once these periods pass, certain things can no longer be learned, (4) teenagers who sleep a lot are just lazy, (5) differences in dominant brain hemispheres (left-brained vs. right-brained) can explain differences in performance and (6) boys have a boy's brain and girls have a girl's brain.

As shown in Figure 1, we included several manipulations throughout the video to influence participants' external on-topic auditory attention. The timing of these manipulations was identical for all participants because, in addition to NET, we also examined brain-to-brain synchrony, and counterbalancing was not feasible for that analysis. For the attention enhancing condition, a red frame appeared around the video together with a big red exclamation mark in the upper left corner. Participants were informed beforehand that they had to pay extra attention when these signs appeared, as they were sure to be asked questions about these parts. For the distracting condition, three types of distractors were used: (1) construction noise added to the audio, (2) an animal video without audio shown in the bottom right corner, and (3) classroom noise added to the audio. In the attentional disengagement condition, participants were instructed to ignore the video lecture while performing an alternative tasks: (1) completing several math exercises (e.g. 16 + 35 = ... + 47 = ... + 89 =... - 28 = ...), (2) reading a text (66 sentences) and (3) counting down in steps of 17 starting from 994 based on Ki and colleagues (2016). At the start of the experiment, participants were given the text and math exercises on paper, along with instructions to complete these tasks when the message appeared on the screen, until they heard a beep. They were given the explicit instruction to focus completely on the alternative tasks and to not listen to the information presented in the video during that time to induce attentional disengagement from the video lecture. Each of the three manipulation types (enhancing, distracting, disengagement) consisted of three manipulation windows of 90 seconds, and each block of the video contained only one type of manipulation. This resulted in a total of 4.5 minutes of data per manipulation. The data in between these manipulations were used as a baseline condition, some of which was used to train the EEG decoding algorithm (see 2.4.1.3. Stimulus reconstruction and correlation).

Block 1 did not contain any manipulations and was entirely used for training the decoding algorithm (see 2.4.1.3. Stimulus reconstruction and correlation). Block 2 consisted of two enhancers, one in the beginning and one at the end. In between these enhancers, there were four minutes that were used for training the decoding algorithm. Block 3 was the same as Block 2 except with two times the disengagement condition instead of the enhancing condition. Block 4 contained no manipulations and the first six minutes were used for training of the decoder. Block 5 contained three distractors and no part that was used for training. Block 6 had one enhancer in the second half of the block, and the first four minutes of the block were used for decoder training. Block 7 was again the same as Block 6 except with the disengagement condition instead of the enhancing condition. After each block, participants answered questions about their external on-topic attention during the past block. They received this self-report questionnaire on paper at the start of the experiment, together with the alternative tasks for the disengagement condition.

The EEG data were collected using a BioSemi Active Two system with 64 electrodes (10-20 system) at a sampling rate of 2048 Hz (*Biosemi,* n.d.). EEG was recorded while the participants watched the video lecture. To synchronise the EEG signal with the video, a flashing square was added to the upper right corner of the video. The square alternated between black and white every minute, which was monitored using a photosensor. This flashing square was not visible to the participants, as the photosensor was taped over the blinking corner on the screen.

Figure 1



Lees de tekst op p.5-6 aandachtig tot je een geluidssignaal hoort. Richt je aandacht volledig op de tekst en NIET meer op de video!

The figure provides a comprehensive overview of our protocol. The video consisted of 7 blocks, which are indicated with different colours, each addressing a different neuromyth. After each part, the participants had to score their own external on-topic attention with a questionnaire and after the video they had to do a performance test. The small (coloured) rectangles represent the nine manipulation windows and the larger (grey) rectangles indicate the parts that were used for training of the decoder. The little circles indicate which item on the performance test belongs to that part of the video.

2.3. Materials

Attention self-report

A self-report questionnaire was administered at fixed time points throughout the video to assess participants' self-reported external on-topic attention levels. As previously discussed, the video was divided into seven blocks, and after each block, participants had to answer five questions regarding their attention levels. These questions were (1) How attentive were you during the previous block?, (2) How well do you think you processed the information from the previous block?, (3) How would you rate your concentration during the previous block?, (4) How long were you distracted from the lesson during the previous block? and (5) To what extent were you able to maintain your focus during the previous block? Responses were recorded on a 10-point Likert scale. To ensure that the participants read the questions carefully, the order of the questions was varied after each block.

For each block, a mean attention score was calculated using the Likert scores on the five questions. The score for the question measuring distraction was reversed so that a higher score indicated more external on-topic attention rather than more distraction.

Performance test

After they watched the video, participants completed a paper-and-pencil performance test consisting of 37 items about information presented in the video lecture, which was included as an indirect measure of external on-topic attention. The items were distributed equally across the four conditions, with seven items pertaining to the attention enhancement condition, seven to the attention distraction condition and seven to the attentional disengagement condition. The remaining 16 items were from the baseline condition without any manipulations.

The test included both open and multiple-choice questions, with each condition having an equal number of open ($n = 3$) and multiple-choice ($n = 4$) questions. For the baseline condition, there were seven open questions and nine multiple-choice questions.

To minimise guessing, participants had the option to indicate that they did not know the answer to a question. Additionally, they could indicate if the information needed to answer the question was not present in the video. This option was included to discern whether participants did not remember the information or did not hear it. All necessary information to answer the questions was included in the video.

Participants received a score of one for each correctly answered question and a score of zero for each incorrectly answered item. Similarly, if participants indicated to not know the answer or to not have heard the information, a score of zero was given. The theoretical maximum score on the performance test was 37. For some analyses, the percentage correct answers per condition was used as a performance measure to compensate for the difference in the number of items between the three conditions and the baseline condition.

2.4. Analyses

2.4.1. NET measure

2.4.1.1. Preprocessing

In this study, we applied two stages of filtering to the EEG signal to prepare it for further analysis. First, we used a first-order IIR (infinite impulse response) high-pass filter with a cutoff frequency of 0.01 Hz to remove very low-frequency components, such as baseline drift and other slow fluctuations. After high-pass filtering, we downsampled the data to a sampling rate of 128 Hz using MATLAB (version 9.11 - 2021b) built-in ‘resample’ command (applying appropriate anti-aliasing filters).

Next, we applied a fourth-order IIR bandpass filter to focus on the key frequency components between 0.1 Hz and 4 Hz (Li et al., 2025). This frequency band was selected because it captures slow brain activities related to attention and cognition, which are crucial for phase-locking to the speech envelope, the slow variations over time of the speech stimulus. After this filtering step, the data was further downsampled to 16 Hz to improve the efficiency of subsequent signal processing algorithms.

2.4.1.2. Speech Envelope Extraction

The speech envelope served as the audio feature in this study. Prior research indicates that EEG signals can phase-lock to the speech envelope (O’Sullivan et al., 2015). To extract this feature, we employed the method from Biesmans et al. (2017), which uses a gammatone filterbank and a power law compression to simulate the frequency selectivity of the human cochlear system.

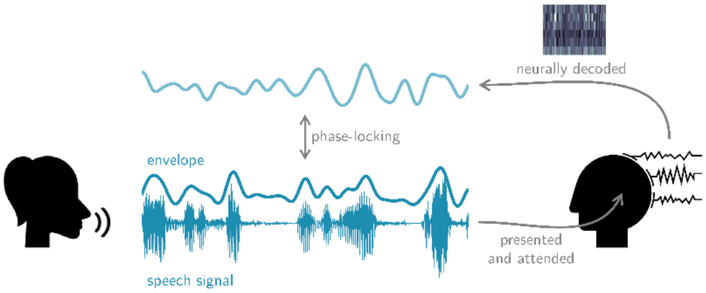
2.4.1.3. Stimulus Reconstruction and Correlation

We adopted a backward model to train a linear spatial-temporal decoder on a per-subject basis using the 27 minutes of data in the grey boxes in Figure 1. The goal of this decoder is to filter out neural activity that is not related to the speech stimulus, and to reshape and combine the stimulus-following neural responses to generate an output that is as close as possible to the original speech envelope. This is achieved by means of a linear regression on nine time-lagged copies of each EEG-channel with the speech envelope as the target and using the least-squares criterion (Biesmans et al., 2017). During the validation phase, we applied the trained decoder on 30s windows and calculated the Pearson correlation coefficient between the reconstructed speech envelope (i.e. the output signal of the decoder) and the ground-truth speech envelope (Figure 2). We hypothesised that a higher correlation value (hence stronger neural responses to the speech stimulus) indicates a higher level of external on-topic attention, i.e. to the teacher's voice (Roebben et al., 2024). In the rest of the paper, we will use the acronym NET (i.e. neural envelope tracking) to refer to the correlation between this reconstructed speech envelope and the ground-truth speech envelope, each time measured over a window of 30 seconds.

In order to further analyse the data, we average all NET values per condition obtained within the manipulation window (as indicated by the coloured boxes in Figure 1), as well as across entire blocks of the video. The latter includes a combination of data in which attention was manipulated and data in which it was not (note that each block contains only one type of manipulation). This last measure, NET averaged across an entire block, is included to enable comparison with behavioural data from the self-reported attention questionnaire, in which participants rated their external on-topic attention over the same time span.

Figure 2

Schematic representation of NET



The figure (adapted from Geirnaert, 2022) illustrates NET, which computes a correlation between the reconstructed speech envelope based on the participants' brain signals and the actual speech stimulus of the teacher.

2.4.2. Brain-to-brain synchrony

To compute the brain-to-brain synchrony, we used the "Maximum Variance" (MAXVAR) Generalised Canonical Correlation Analysis (MAXVAR-GCCA) framework (Carroll, 1968; Horst, 1961; Kettenring, 1971). This method computes a set of subject-specific spatio-temporal filters that project each participant's EEG into a shared low-dimensional subspace, such that the projected signals across subjects are as close as possible.

The preprocessing for the ISC analysis followed the same overall structure as the general EEG pipeline. Slow drifts were removed using a first-order IIR high-pass filter at 0.5 Hz. The EEG was then downsampled to 128 Hz, after which a fourth-order IIR band-pass filter between 1 and 4 Hz was applied. Following this filtering step, the data were further downsampled to 16 Hz.

MAXVAR-GCCA filters were trained on the designated 27-minutes training data (see grey boxes Figure 1). For each subject, a spatio-temporal filter was learned to map the multichannel EEG to components capturing the neural activity most consistently shared across the group. We retained the first six GCCA components in all analyses.

ISC was computed separately for each time window. For a given subject, that subject's GCCA components were correlated with the corresponding components from every other participant (29 subject pairs in total). Correlations were computed component-wise, and the six component-wise correlations were summed to obtain the ISC score for that subject in that window. Higher ISC values indicate stronger alignment of a participant's neural response with the group.

2.4.3. Alpha power

The EEG data were pre-processed using a high-pass filtering step. Specifically, a first-order IIR (infinite impulse response) high-pass filter with a cutoff frequency of 0.01 Hz was applied to remove very low-frequency components, including baseline drift and other slow fluctuations. The filtered signal was then downsampled to 128 Hz using MATLAB (version 9.11 – 2021b) with the built-in resample function, which incorporates appropriate anti-aliasing filters.

To quantify alpha power, the MATLAB (version 9.11 – 2021b) bandpower function was used to compute the spectral power in the 8–12 Hz range (alpha band) for each of the 64 EEG channels. This analysis was performed using 30-second windows, resulting in 64 alpha power values per window. These values were then averaged to obtain a single global alpha power measure for each 30-second window.

2.4.4. Statistical Analysis of Behavioural and EEG Data

The behavioural data and the EEG data were further analysed using JASP (Version 0.19.0) and rmcorr-shiny (Bakdash & Marusich, 2017; Marusich & Bakdash, 2021). To analyse the effect of our manipulations on self-reported attention, performance test scores, NET, brain-to-brain synchrony and alpha power, we used (generalised) linear mixed-effects models, given that our data was nested within participants. A generalised mixed-effects model analysis were employed for the performance test scores, given that the responses were either correct (1) or incorrect (0). For all other measures, a linear mixed-effects model was used. We examined whether the accuracy on the performance test, the self-reported attention, NET, brain-to-brain synchrony and alpha power differed between the conditions. For NET we did this analysis when these were measured during the manipulation windows and when these were measured over an entire block.

Furthermore, we aimed to investigate the association between the behavioural measures and the EEG measures of external on-topic attention. On the one hand, we looked at the correlation between the self-reported attention scores and the corresponding NET across the entire block combined per condition. Specifically, this means that for every participant we had an average self-reported attention score and NET measured over an entire block for every condition. On the other hand, we looked at the correlation between the performance test scores and the corresponding NET measured during the manipulation window for every condition. Again, we had for each participant a percentage correct answers linked with NET from the respective manipulation window. Therefore we calculated a repeated measures correlation, which is a technique to calculate the intra-individual association between two paired measures while accounting for the non-independence of the data (Bakdash & Marusich, 2017). The analyses were conducted using rmcorr-shiny (Bakdash & Marusich, 2017; Marusich & Bakdash, 2021) . An alpha level of .05 (two-tailed) was used to determine statistical significance in all analyses, and multiple comparisons were corrected using the Holm method.

3. Results

3.1. Behavioural results

The data for the self-reported attention, separated by condition, are shown in Figure 3a. A linear mixed-effects model revealed a significant effect of manipulation on the self-reported attention scores [$F(3,177) = 51.90$, $p < .001$]. Further Holm-corrected contrast analyses revealed significant differences between all conditions, except between baseline and enhancing. These findings indicated that both the distractors and the alternative tasks in the disengagement condition resulted in lower self-reported external on-topic attention, but participants did not report higher external on-topic attention in the enhancing condition (Table 1).

Similar results were observed for the performance test scores (Figure 3b). A generalised linear mixed-effects model confirmed a significant effect of manipulation on performance [$\chi^2(3) = 262.32$, $p < .001$]. Contrast analyses revealed significant differences between all manipulations in the expected direction after Holm correction (Table 1). Distraction and attentional disengagement lowered the performance, while enhancing improved it. Although participants did not report higher external on-topic attention under the enhancing condition, these results suggest that they did remember more from these periods.

Table 1

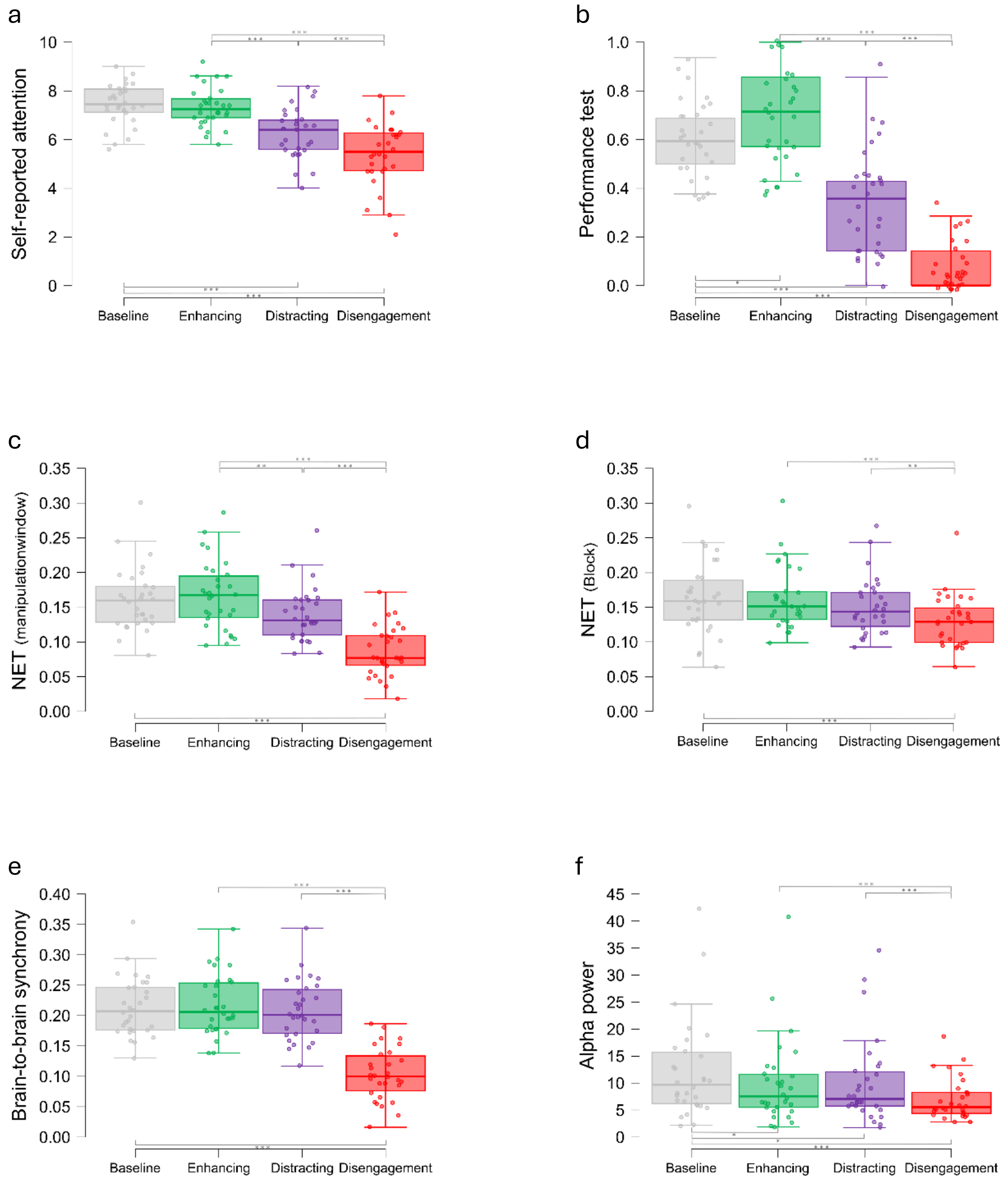
Contrast analyses for behavioural measures

		Estimate	SE	z	p_{holm}
		Self-reported attention scores			
Baseline	Enhancing	0.13	0.19	0.69	.491
	Distracting	1.21	0.23	5.23	< .001***
	Disengagement	2.06	0.19	10.94	< .001***
Enhancing	Distracting	1.08	0.23	4.67	< .001***
	Disengagement	1.93	0.19	10.25	< .001***
Distracting	Disengagement	0.86	0.23	3.70	< .001***
		Performance test			
Baseline	Enhancing	-0.09	0.04	-2.26	.024*
	Distracting	0.26	0.04	6.46	< .001***
	Disengagement	0.55	0.03	17.43	< .001***
Enhancing	Distracting	0.35	0.05	7.54	< .001***
	Disengagement	0.64	0.04	16.69	< .001***
Distracting	Disengagement	0.29	0.04	7.24	< .001***

*$p < .05$. **$p < .01$. ***$p < .001$.

Figure 3

Behavioural and EEG measures for each condition



*The left panel above (3a) shows the self-reported attention scores; the right panel above (3b) shows the proportion correct answers on the performance test; the left panel in the middle (3c) shows NET during the manipulation windows; the right panel in the middle (3d) shows NET for an entire block for each condition, with block 4 as baseline, blocks 2 and 6 as the enhancing condition, blocks 3 and 7 as the disengagement condition and block 5 as the distracting condition; the left panel below (3e) shows brain-to-brain synchrony during the manipulation window and the right panel below (3f) shows alpha power during the manipulation window. Each dot in these plots represents the mean per condition of a single participant. The gray lines indicate whether the difference between the conditions was significant (*p<.05. **p<.01. ***p<.001.).*

3.2. NET

Figure 3c shows NET during the manipulation windows, separated by condition. A linear mixed-effects model revealed a significant effect of manipulation on NET [$F(3,267) = 39.29$, $p < .001$]. The results of the contrast analyses using the Holm correction are presented in Table 2. NET was significantly lower in the disengagement condition compared to all other conditions and in the distracting condition compared to the enhancing condition. No significant differences were found between baseline and either enhancing or distracting.

Next, we examined NET measured over an entire block (Figure 3d), which consisted of a mix of data where attention was manipulated and data where no manipulation occurred. A linear mixed-effect model analysis again found a significant effect of manipulation on NET [$F(3,147) = 18.16$, $p < .001$]. The Holm-adjusted contrasts (Table 2) revealed that only the disengagement condition had significantly lower NET compared to all other conditions.

Table 2

Contrast analyses for NET

		Estimate	SE	z	p_{holm}
		Manipulation windows			
Baseline	Enhancing	-0.01	0.01	-0.51	.611
	Distracting	0.02	0.01	1.98	.095
	Disengagement	0.07	0.01	6.75	<.001***
Enhancing	Distracting	0.03	0.01	3.52	.001**
	Disengagement	0.08	0.01	10.27	<.001***
Distracting	Disengagement	0.05	0.01	6.74	<.001***
		Entire block			
Baseline	Enhancing	$-6{,}72\text{x}10^{-5}$	0.01	-0.01	.991
	Distracting	0.01	0.01	1.62	.211
	Disengagement	0.03	0.01	5.49	<.001***
Enhancing	Distracting	0.01	0.01	1.88	.180
	Disengagement	0.03	0.01	6.74	<.001***
Distracting	Disengagement	0.02	0.01	3.62	.001**

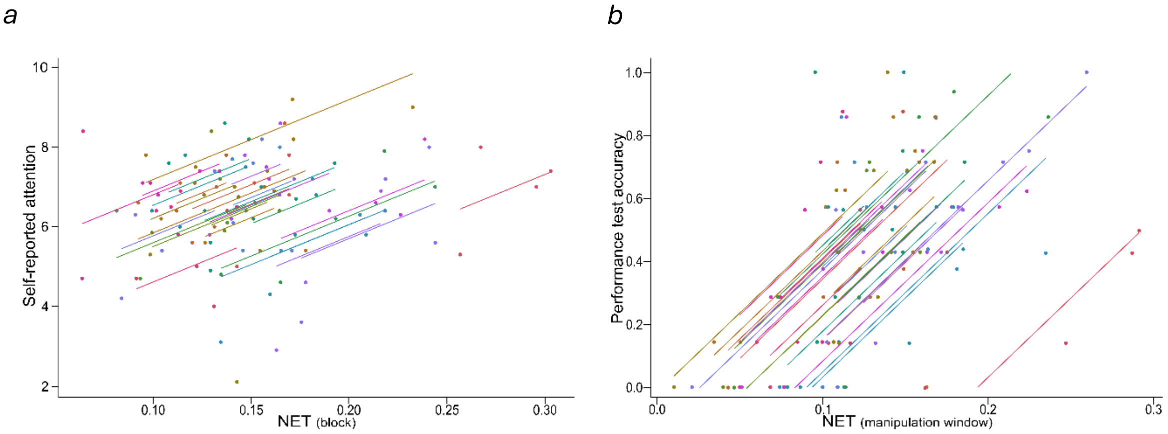
*$p < .05$. **$p < .01$. ***$p < .001$.

3.3. Associations between behavioural measures and NET

We further examined the association between the behavioural data and NET via a repeated measures correlation analysis. The findings reveal a moderate repeated measure correlation between NET measured over an entire block and self-reported attention scores ($r_{rm} = .41$; $p < .001$; Figure 4a). This means that an increase in NET, which reflects higher external on-topic auditory attention, is associated with higher self-reported attention. A strong correlation was observed between NET measured exactly during the manipulation window and performance on the corresponding items ($r_{rm} = .69$; $p < .001$; Figure 4b).

Figure 4

Correlations between behavioural data and EEG data



The left panel (4a) illustrates the correlations between self-reported attention and NET measured over an entire block. The right panel (4b) illustrates the correlation between the performance test accuracy and NET during the manipulation windows. The different colours represent individual participants, with each participant having four dots that indicate their average scores for each condition. The slope of the lines reflects the overall correlation between the behavioural an the EEG data.

3.4. Brain-to-brain synchrony

Brain-to-brain synchrony (Figure 3e) during the manipulation window differed significantly across conditions [$F(3,267) = 129.41$, $p < .001$]. A contrast analysis indicated that brain-to-brain synchrony was only significantly lower in the disengagement condition compared to the other conditions after Holm correction, whereas baseline, enhancing and distracting conditions showed very similar synchrony levels (Table 3).

3.5. Alpha power

Similarly, alpha power (Figure 3f) during the manipulation window was significantly different across conditions [$F(3,267) = 19.25$, $p < .001$]. Contrast analyses using Holm correction revealed that alpha power was significantly higher in baseline compared to the other conditions. Alpha power was lower in the disengagement condition, as compared to the enhancing and distraction conditions, while the latter two did not differ (Table 3).

Table 3

Contrast analyses for brain-to-brain synchrony and alpha power

		Estimate	SE	z	p_{holm}
		Brain-to-brain synchrony			
Baseline	Enhancing	-0.00	0.01	-0.38	1.00
	Distracting	0.01	0.01	0.55	1.00
	Disengagement	0.11	0.01	11.82	<.001***
Enhancing	Distracting	0.01	0.01	1.32	.564
	Disengagement	0.11	0.01	17.26	<.001***
Distracting	Disengagement	0.11	0.01	15.94	<.001***
		Alpa power			
Baseline	Enhancing	2.13	0.77	2.76	.018*
	Distracting	2.02	0.77	2.61	.018*
	Disengagement	5.06	0.77	6.51	<.001***
Enhancing	Distracting	- 0.11	0.55	- 0.20	.842
	Disengagement	2.90	0.55	5.30	<.001***
Distracting	Disengagement	3.01	0.55	5.50	<.001***

**p* < .05. ***p* < .01. ****p* < .001.

4. Discussion

This study examined the potential of neural envelope tracking (NET) as an indicator of external on-topic auditory attention during learning. NET reflects the strength of the relation between the EEG-based reconstructed speech stimulus (envelope) and the actual speech stimulus (envelope) (Biesmans et al., 2017; Roebben et al., 2024; Vanthornhout et al., 2019). To examine its validity, we introduced attention manipulations (distractors and enhancers) during a lecture. The findings indicated that NET measured during the manipulation window differentiated between most experimental conditions, with the exception that baseline did not differ from enhancing or distracting. To further validate NET, we included a performance test, as an indirect measure of attention, and a self-reported attention questionnaire. Both methods effectively distinguished between nearly all conditions, thereby confirming that our manipulations effectively modulated the participants' external on-topic attention. Performance and self-reported attention were lower in the distracting and disengagement condition compared to baseline. The enhancing condition led to improved performance, but did not result in higher self-reported attention. This discrepancy may stem from the fact that self-reported attention scores reflected attention during an entire block, including only a small part of enhanced attention, whereas performance test scores were based on information from the manipulation windows. NET measured during the manipulation window strongly correlated with performance, while NET measured over an entire block demonstrated a weaker yet still statistically significant association with self-reported attention. These findings suggest that NET is sensitive to fluctuations in external on-topic auditory attention during learning.

A consistent trend that emerged is that NET measured during the manipulation window was lower during the disengagement condition compared to all other conditions. This suggests that NET decreases during moments where the attention shifts to something else than the lecture. A similar effect was observed by Roebben et al. (2024) and Vanthornhout et al. (2019), who investigated NET using children's stories by including alternative tasks into their paradigms and also observed reduced NET in these attentional disengagement conditions. Our findings extend these findings by applying it in an educational context using longer stimuli, highlighting its potential as an indicator for external on-topic auditory attention to an educational stimulus.

For the distracting condition, in which participants were distracted by background noise or a silent video, there was only a significant difference with the enhancing condition when NET was measured during the manipulation window. No significant difference was observed between baseline and the distracting condition. However, because NET was slightly higher in the enhancing condition and slightly lower in the distracting condition compared to baseline when measured during the manipulation window, the difference between enhancing and distracting was significant. This finding aligns with the observations reported by Levy, Korisky, et al. (2025), who examined the impact of background noise on attention decoding and similarly found slightly, but not significantly, reduced NET in the noise conditions. While in the study by Levy, Shadi, et al. (2025) NET was significantly lower in the intermittent noise and quiet condition compared to the continuous noise condition. Taken together, background noise does not uniformly lower NET, while one study reported increased NET under continuous noise (Levy, Shadi, et al., 2025), our data and another study (Levy, Korisky, et al., 2025) showed small, non-significant reductions, underscoring the need to determine how noise type and contextual factors modulate this effect. In the enhancing condition NET did not increase enough to be significantly different from baseline. To the best of our knowledge, no previous studies have examined the effect of attention enhancement on NET. This underscores the necessity for future studies to investigate the effect of attention enhancement on NET.

Besides NET measured during the manipulation window, we also measured NET over an entire block, to facilitate a fairer comparison with the self-reported attention, which required participants to rate their attention over the same period. This measure included both manipulated and non-manipulated data, likely diluting the effects, as no differences were observed between baseline, enhancing and distracting conditions. Averaging NET across an entire block reduced its precision, although it is one of the key strengths of EEG to track cognitive processes during learning rather than retrospectively. However, this step was important to link NET with the self-reported attention. Future studies should use NET for continuous monitoring of attention during learning, which will be more sensitive to changes in external on-topic auditory attention, similarly as our NET measured during the manipulation window.

To compare NET with more traditional neural markers of attention during learning, we examined both brain-to-brain synchrony and alpha power. However, neither measure reliably indicated when participants were attentive to the teacher. The key difference between brain-to-brain synchrony and NET was that the latter did not reflect reduced attention during distraction, unlike NET and behavioural data. This likely occurred because distractions still produced a shared experience: participants were exposed to the same distracting stimuli at the same time, maintaining synchrony. In contrast, during the disengagement condition, synchrony decreased even though participants performed the same alternative task. Individual differences in reading and calculation speed likely reduced the shared experience, explaining this drop in synchrony. Furthermore, using brain-to-brain synchrony has an additional drawback that it can only be measured if EEG recordings of multiple students listening to the same lecturer are available. The main difference between alpha power and NET appeared in the attentional disengagement condition, where alpha power was lowest, suggesting the highest attention, while both NET and behavioural measures indicated low attention for this condition. This discrepancy is not unexpected, as participants were not inattentive per se, but were engaged in the alternative task rather than the lecture. Thus, both brain-to-brain synchrony and alpha power effectively capture overall attentional engagement but cannot distinguish whether attention is directed toward the lecture or somewhere else. Consequently, these measures may indicate heightened attention even when attention is not focused on the intended learning material.

This distinction explains why other studies successfully used both brain-to-brain synchrony and alpha power to measure classroom attention (Bevilacqua et al., 2019; Cohen et al., 2018; Davidesco et al., 2023; Dikker et al. 2017; Dikker et al., 2020; Grammer et al., 2021). Those studies examined attention across different teaching styles, time of the day, or testing effects, assuming students remained focused on the educational content in all conditions. In contrast, this study explicitly manipulated external on-topic attention using distractors and alternative tasks.

Overall, the results from both more traditional measures, brain-to-brain synchrony and alpha power, highlight the importance of taking the stimulus into account as NET does, rather than simply measuring overall attentional engagement. Results from brain-to-brain synchrony and alpha power were difficult to reconcile with behavioural data, as both indicated higher attention in conditions designed to reduce it (distracting condition for brain-to-brain synchrony and disengagement condition for alpha power). In contrast, NET aligned more closely with the behavioural measures while providing objective and continuous insight into external on-topic attention that behavioural measures alone cannot capture, suggesting that NET may serve as a potential marker of external on-topic attention in individual learners.

4.1. Limitations and future directions

While the present study provides valuable insights into NET as potential indicator of external on-topic auditory attention for an educational stimulus, several limitations should be considered when interpreting the findings. First, to compare NET with more traditional measures of attention during learning, we included brain-to-brain synchrony and alpha power. Calculating brain-to-brain synchrony required all participants to complete the exact same experiment, which prevented counterbalancing conditions across participants. As a result, we could not fully control for order effects such as fatigue. Future studies should counterbalance conditions to ensure that observed differences in NET can be more confidently attributed to the experimental manipulations.

Second, although we made the lecture as educationally realistic as possible by using real lecture content, participants were seated alone in the lab, which differs from a real classroom setting. We also used a 64-electrode EEG system with wet electrodes, which is impractical to use in real classroom settings or to simultaneously collect EEG data in multiple participants. Future research should therefore consider the use of user-friendly mobile EEG devices. However, these systems typically have a lower number of electrodes, and tend to be more sensitive to various artifacts, leading to increased data loss, making further improvements necessary for broader implementation in educational research (Janssen et al., 2021). Despite these challenges, mobile-EEG systems would allow to use NET in a real classroom environment, with all its inherent features, such as classmates, a familiar teacher, and more real-world distractions. This is important because previous studies that used brain-to-brain synchrony to measure attention have shown that this measure is influenced by factors such as social closeness between classmates (Dikker et al., 2020) and between students and teachers (Bevilacqua et al., 2019), but it remains unclear whether such factors also impact NET and this represents an avenue for future study.

Third, the baseline condition was defined as a condition without experimental manipulation of external on-topic attention. Nevertheless, naturally occurring attentional fluctuations are unavoidable, due to factors such as mind-wandering, fatigue and environmental distractions. Treating baseline as a stable reference in comparison with other conditions may bias observed differences, making them appear larger or smaller than they truly are. Future research should therefore aim to construct a more reliable baseline reference condition by excluding periods associated with attentional lapses using more fine-grained methods such as a self-caught method with clickers (Bunce et al. 2010; Hlas et al., 2009).

Lastly, training the decoder required a large amount of baseline data. This necessitated extended baseline recordings, increasing participant demands and limiting the feasibility of implementation in real classroom environments. Consequently, the overall experiment duration was longer, but nearly half of the EEG data was used for training of the decoder, reducing the data available for analysis. Future research should aim to reduce the amount of training data required, for example, by starting from universal decoders and applying unsupervised (time-adaptive) training to gradually finetune them towards a personalised decoder (Geirnaert et al., 2022; Heintz et al., 2025). This would decrease, or even remove, the need to explicitly define and collect baseline data for training the decoder and potentially enable real-time feedback on students' external on-topic attention in the future.

4.2. Conclusion

The current study offers preliminary evidence that neural envelop tracking may serve as a potential indicator of external on-topic auditory attention to an educational stimulus. By incorporating the instructor’s voice into the calculation of the attention measure, neural envelope tracking could provide a more accurate indicator of external on-topic attention to the teacher compared to more general indicators of attention, such as brain-to-brain synchrony and alpha power. Future research should optimise this technique by focusing on more user-friendly EEG devices, such as mobile EEG, the use of un-supervised and/or time-adaptive decoders, and by studying its applications in more diverse populations and educational settings. This will allow us to explore the potential of NET into real-world educational studies.

References

Anderson, B. (2011). There is no such thing as attention. *Frontiers in Psychology, 2*, Article 246. https://doi.org/10.3389/fpsyg.2011.00246

Bakdash, J. Z., & Marusich, L. R. (2017). Repeated Measures Correlation. *Frontiers in Psychology*, *8,* Article 456. https://doi.org/10.3389/fpsyg.2017.00456

Bevilacqua, D., Davidesco, I., Wan, L., Chaloner, K., Rowland, J., Ding, M., Poeppel, D., & Dikker, S. (2019). Brain-to-brain synchrony and learning outcomes vary by student–teacher dynamics: Evidence from a real-world classroom electroencephalography study. *Journal of Cognitive Neuroscience*, *31*(3), 401–411. https://doi.org/10.1162/jocn_a_01274

Biesmans, W., Das, N., Francart, T., & Bertrand, A. (2017). Auditory-inspired speech envelope extraction methods for improved EEG-based auditory attention detection in a cocktail party scenario. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *25*(5), 402–412. https://doi.org/10.1109/TNSRE.2016.2571900

Biosemi. (n.d.). *Active2 and Active 2.5: 280-channel, DC amplifier, 24-bit resolution, biopotential measurement system with active electrodes.* https://www.biosemi.com/Products_ActiveTwo.htm

Blasiman, R. N., Larabee, D., & Fabry, D. (2018). Distracted students: A comparison of multiple types of distractions on learning in online lectures. *Scholarship of Teaching and Learning in Psychology*, *4*(4), 222–230. https://doi.org/10.1037/stl0000122

Booth, B. M., Bosch, N., & D'Mello, S. K. (2023). Engagement detection and its applications in learning: A tutorial and selective review. *Proceedings of the IEEE*, *111*(10), 1398–1422. https://doi.org/10.1109/JPROC.2023.3309560

Boudewyn, M. A., & Carter, C. S. (2018). I must have missed that: Alpha-band oscillations track attention to spoken language. *Neuropsychologia*, *117*, 148–155. https://doi.org/10.1016/j.neuropsychologia.2018.05.024

Bradbury, N. A. (2016). Attention span during lectures: 8 seconds, 10 minutes, or more? *Advances in Physiology Education*, *40*(4), 509–513. https://doi.org/10.1152/advan.00109.2016

Bühler, B., Bozkir, E., Goldberg, P., Deininger, H., D'Mello, S., Gerjets, P., Trautwein, U., & Kasneci, E. (2025). Temporal dynamics of meta-awareness of mind wandering during lecture viewing: Implications for learning and automated assessment using machine learning. *Journal of Educational Psychology*, *117*(1), 38–62. https://doi.org/10.1037/edu0000903

Bunce, D. M., Flens, E. A., & Neiles, K. Y. (2010). How long can students pay attention in class? A study of student attention decline using clickers. *Journal of Chemical Education*, *87*(12), 1438–1443. https://doi.org/10.1021/ed100409p

Carroll, J. D. (1968). Generalization of canonical correlation analysis to three of more sets of variables. *Proceedings of the 76th annual convention of the American Psychological Association, 3, 227-228.*

Cohen, R. A. (2014). *The neuropsychology of attention*. Springer. https://doi.org/10.1007/978-0-387-72639-7

Cohen, S. S., Madsen, J., Touchan, G., Robles, D., Lima, S. F. A., Henin, S., & Parra, L. C. (2018). Neural engagement with online educational videos predicts learning performance for individual students. *Neurobiology of Learning and Memory*, *155*, 60–64. https://doi.org/10.1016/j.nlm.2018.06.011

Compton, R. J., Gearinger, D., & Wild, H. (2019). The wandering mind oscillates: EEG alpha power is enhanced during moments of mind-wandering. *Cognitive, Affective, & Behavioral Neuroscience*, *19*(5), 1184–1191. https://doi.org/10.3758/s13415-019-00745-9

Davidesco, I. (2020). Brain-to-brain synchrony in the STEM classroom. *CBE* Life Sciences Education, *19*(3). Article es8. https://doi.org/10.1187/cbe.19-11-0258

Davidesco, I., Laurent, E., Valk, H., West, T., Milne, C., Poeppel, D., & Dikker, S. (2023). The temporal dynamics of brain-to-brain synchrony between students and teachers predict learning outcomes. *Psychological Science*, *34*(5), 633–643. https://doi.org/10.1177/09567976231163872

De Smedt, B. (2018). Applications of cognitive neuroscience in educational research. In *Oxford Research Encyclopedia of Education*. https://doi.org/10.1093/acrefore/9780190264093.013.69

Dikker, S., Haegens, S., Bevilacqua, D., Davidesco, I., Wan, L., Kaggen, L., McClintock, J., Chaloner, K., Ding, M., West, T., & Poeppel, D. (2020). Morning brain: Real-world neural evidence that high school class times matter. *Social Cognitive and Affective Neuroscience*, *15*(11), 1193–1202. https://doi.org/10.1093/scan/nsaa142

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., Rowland, J., Michalareas, G., Bavel, J. J. V., Ding, M., & Poeppel, D. (2017). Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Current Biology*, *27*(9), 1375–1380. https://doi.org/10.1016/j.cub.2017.04.002

Di Lollo, V. (2018). Attention is a sterile concept; iterative reentry is a fertile substitute. *Consciousness and Cognition, 64*, 45–49. https://doi.org/10.1016/j.concog.2018.02.005

Dmochowski, J. P., Sajda, P., Dias, J., & Parra, L. C. (2012). Correlated components of ongoing EEG point to emotionally laden attention: A possible marker of engagement? *Frontiers in Human Neuroscience*, *6*, Article 112. https://doi.org/10.3389/fnhum.2012.00112

Forrin, N. D., Huynh, A. C., Smith, A. C., Cyr, E. N., McLean, D. B., Siklos-Whillans, J., Risko, E. F., Smilek, D., & MacLeod, C. M. (2021). Attention spreads between students in a learning environment. *Journal of Experimental Psychology: Applied*, *27*(2), 276–291. https://doi.org/10.1037/xap0000341

Foxe, J., & Snyder, A. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in Psychology*, *2,* Article 154. https://doi.org/10.3389/fpsyg.2011.00154

Geirnaert, S. (2022) Single processing algorithms for EEG-based auditory attention decoding [Doctoral dissertation, KU Leuven] Lirias. https://lirias.kuleuven.be/retrieve/662156

Geirnaert, S., Francart, T., & Bertrand, A. (2022). Time-adaptive unsupervised auditory attention decoding using EEG-based stimulus reconstruction. *IEEE Journal of Biomedical and Health Informatics*, *26*(8), 3767–3778. https://doi.org/10.1109/JBHI.2022.3162760

Grammer, J. K., Xu, K., & Lenartowicz, A. (2021). Effects of context on the neural correlates of attention in a college classroom. *Npj Science of Learning*, *6*, Article 15. https://doi.org/10.1038/s41539-021-00094-8

Heintz, N., Francart, T., & Bertrand, A. (2025). *Unsupervised EEG-based decoding of absolute attention with canonical correlation analysis*. ArXiv. https://doi.org/10.48550/arXiv.2504.17724

Hlas, A. C., Neyers, K., & Molitor, S. (2019). Measuring student attention in the second language classroom. *Language Teaching Research*, *23*(1), 107–125. https://doi.org/10.1177/1362168817713766

Hobbiss, M. H., & Lavie, N. (2024). Sustained selective attention in adolescence: Cognitive development and predictors of distractibility at school. *Journal of Experimental Child Psychology*, *238*, Article 105784. https://doi.org/10.1016/j.jecp.2023.105784

Hommel, B., Chapman, C. S., Cisek, P., Neyedli, H. F., Song, J.-H., & Welsh, T. N. (2019). No one knows what attention is. *Attention, Perception, & Psychophysics*, *81*(7), 2288–2303. https://doi.org/10.3758/s13414-019-01846-w

Horst, P. (1961). Generalized canonical correlations and their application to experimental data. *Journal of Clinical Psychology, 17*(4), 331-347. https://doi.org/10.1002/1097-4679(196110)17:4%3C331::aid-jclp2270170402%3E3.0.co;2-d

Howard-Jones, P. (2010). *Introducing neuroeducational research: Neuroscience, education and the brain from contexts to practice*. Routledge.

Janssen, T. W. P., Grammer, J. K., Bleichner, M. G., Bulgarelli, C., Davidesco, I., Dikker, S., Jasińska, K. K., Siugzdaite, R., Vassena, E., Vatakis, A., Zion-Golumbic, E., & van Atteveldt, N. (2021). Opportunities and limitations of mobile neuroimaging technologies in educational neuroscience. *Mind, Brain, and Education*, *15*(4), 354–370. https://doi.org/10.1111/mbe.12302

Jarrett, C. (2015). *Great myths of the brain*. Wiley Blackwell.

Keller, A. S., Davidesco, I., & Tanner, K. D. (2020). Attention matters: How orchestrating attention may relate to classroom learning. *CBE Life Sciences Education*, *19*(3), Article fe5. https://doi.org/10.1187/cbe.20-05-0106

Kettenring, J. R. (1971). Canonical analysis of several sets of variables. *Biometrika, 58*(3), 433-451. https://doi.org/10.2307/2334380

Ki, J. J., Kelly, S. P., & Parra, L. C. (2016). Attention strongly modulates reliability of neural responses to naturalistic narrative stimuli. *The Journal of Neuroscience*, *36*(10), 3092–3101. https://doi.org/10.1523/JNEUROSCI.2942-15.2016

Levy, O., Hackmon, S. L., Zvilichovsky, Y., Korisky, A., Bidet-Caulet, A., Schweitzer, J. B., & Golumbic, E. Z. (2025). Selective attention and sensitivity to auditory disturbances in a virtually real classroom. *eLife, 13*,Article RP103235 https://doi.org/10.7554/eLife.103235.

Levy, O., Korisky, A., Zvilichovsky, Y., & Zion Golumbic, E. (2025). The neurophysiological costs of learning in a noisy classroom: An ecological virtual reality study. *Journal of Cognitive Neuroscience*, *37*(2), 300–316. https://doi.org/10.1162/jocn_a_02249

Levy, O., Shadi, T., Korisky, A., Bleichner, M. G., & Zion Golumbic, E. (2025). Differential effects of external noise and situational interest on neurophysiological responses during video based learning. *Npj Science of Learning*, *10*(1), Article 92. https://doi.org/10.1038/s41539-025-00392-5

Li, T., Geirnaert, S., Van den Broek, D., Bellon, E., De Smedt, B., & Bertrand, A. (2025). *Temporal variation in the acoustic dynamic range is a confounding factor in EEG-based tracking of absolute auditory attention to speech*. Proc. International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Denmark.

Liu, Q., Lin, Y., & Zhang, W. (2025). Intersubject correlation as a predictor of attention: A systematic review. *BMC Psychology, 13,* Article 546. https://doi.org/10.1186/s40359-025-02879-7

Madsen, J., & Parra, L. C. (2022). Cognitive processing of a common stimulus synchronizes brains, hearts, and eyes. *PNAS Nexus, 1,* 1-14. https://doi.org/10.1093/pnasnexus/pgac020

Malik, A. S., & Amin, H. U. (2017). *Designing EEG experiments for studying the brain*. Academic Press. https://doi.org/10.1016/B978-0-12-811140-6.00001-1

Marusich, L. R., & Bakdash, J. Z. (2021). RmcorrShiny: A web and standalone application for repeated measures correlation. *F1000Research, 10,* Article 697. https://doi.org/10.12688/f1000research.55027.1

Matsuo, M., Higuchi, T., Ichibakase, T., Suyama, H., Takahara, R., & Nakamura, M. (2024). Differences in electroencephalography power levels between poor and good performance in attentional tasks. *Brain Sciences*, *14*(6), Article 6. https://doi.org/10.3390/brainsci14060527

Mayer, R. E. (2017). How can brain research inform academic learning and instruction? *Educational Psychology Review*, *29*(4), 835–846. https://doi.org/10.1007/s10648-016-9391-1

Mirkovic, B., Debener, S., Jaeger, M., & De Vos, M. (2015). Decoding the attended speech stream with multi-channel EEG: Implications for online, daily-life applications. *Journal of Neural Engineering*, *12*(4), Article 046007. https://doi.org/10.1088/1741-2560/12/4/046007

O'Connell, R. G., Dockree, P. M., Robertson, I. H., Bellgrove, M. A., Foxe, J. J., & Kelly, S. P. (2009). Uncovering the neural signature of lapsing attention: Electrophysiological signals predict errors up to 20 s before they occur. *Journal of Neuroscience*, *29*(26), 8604–8611. https://doi.org/10.1523/JNEUROSCI.5967-08.2009

Olney, A. M., Risko, E. F., D'Mello, S. K., & Graesser, A. C. (2015). Attention in educational contexts: The role of the learning task in guiding attention. In J. Fawcett, E. Risko, & A. Kingstone (Eds.) *The handbook of attention* (pp. 623-641). The MIT Press. https://doi.org/10.7551/mitpress/10033.003.0030

O'Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinn-Cunningham, B. G., Slaney, M., Shamma, S. A., & Lalor, E. C. (2015). Attentional selection in a cocktail party environment can be decoded from single-trial EEG. *Cerebral Cortex*, *25*(7), 1697–1706. https://doi.org/10.1093/cercor/bht355

Poulsen, A. T., Kamronn, S., Dmochowski, J., Parra, L. C., & Hansen, L. K. (2017). EEG in the classroom: Synchronised neural recordings during video presentation. *Scientific Reports*, *7*, Article 43916. https://doi.org/10.1038/srep43916

Rapport, M. D., Kofler, M. J., Alderson, R. M., Timko, T. M., & DuPaul, G. J. (2009). Variability of attention processes in ADHD: Observations from the classroom. *Journal of Attention Disorders*, *12*(6), 563–573. https://doi.org/10.1177/1087054708322990

Risko, E. F., Buchanan, D., Medimorec, S., & Kingstone, A. (2013). Everyday attention: Mind wandering and computer use during lectures. *Computers & Education*, *68*, 275–283. https://doi.org/10.1016/j.compedu.2013.05.001

R Core Team (2024). _*R: A Language and Environment for Statistical Computing*_ (Version 4.4.1) [Computer software]. https://www.R-project.org/

Roebben, A., Heintz, N., Geirnaert, S., Francart, T., & Bertrand, A. (2024). 'Are you even listening?': EEG-based decoding of absolute auditory attention to natural speech. *Journal of Neural Engineering*, *21*(3), Article 036046. https://doi.org/10.1088/1741-2552/ad5403

Smallwood, J., McSpadden, M., & Schooler, J. W. (2008). When attention matters: The curious incident of the wandering mind. *Memory & Cognition*, *36*(6), 1144–1150. https://doi.org/10.3758/MC.36.6.1144

Straetmans, L., Adiloglu, K., & Debener, S. (2024). Neural speech tracking and auditory attention decoding in everyday life. *Frontiers in Human Neuroscience*, *18*, Article 1483024 https://doi.org/10.3389/fnhum.2024.1483024

Styles, E. A. (2006). *The psychology of attention* (2nd ed.). Psychology Press.

van Atteveldt, N., van Kesteren, M. T. R., Braams, B., & Krabbendam, L. (2018). Neuroimaging of learning and development: Improving ecological validity. *Frontline Learning Research*, *6*(3), 186–203. https://doi.org/10.14786/flr.v6i3.366

Vanthornhout, J., Decruy, L., & Francart, T. (2019). Effect of task and attention on neural tracking of speech. *Frontiers in Neuroscience*, *13*. Article 977. https://doi.org/10.3389/fnins.2019.00977

Varao-Sousa, T. L., & Kingstone, A. (2019). Are mind wandering rates an artifact of the probe-caught method? Using self-caught mind wandering in the classroom to test, and reject, this possibility. *Behavior Research Methods*, *51*(1), 235–242. https://doi.org/10.3758/s13428-018-1073-0

Zeamer, C., & Fox Tree, J. E. (2013). The process of auditory distraction: Disrupted attention and impaired recall in a simulated lecture environment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(5), 1463–1472. https://doi.org/10.1037/a0032190