

This is a repository copy of *An automatic sustained attention prediction (ASAP) method for infants and toddlers using wearable device signals*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/225011/>

Version: Published Version

Article:

Zhang, Yisi, Martinez Cedillo, Priscilla orcid.org/0000-0001-7327-3614, Mason, Harry Thomas orcid.org/0000-0002-3464-7254 et al. (5 more authors) (2025) An automatic sustained attention prediction (ASAP) method for infants and toddlers using wearable device signals. Scientific Reports. 13298 (2025). ISSN 2045-2322

<https://doi.org/10.1038/s41598-025-96794-x>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:



<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



OPEN An automatic sustained attention prediction (ASAP) method for infants and toddlers using wearable device signals

Yisi Zhang^{1,10}, A. Priscilla Martinez-Cedillo^{2,3,10}, Harry T. Mason^{4,5,10}, Quoc C. Vuong^{6,7}, M. Carmen Garcia-de-Soria^{2,8}, David Mullineaux⁹, Marina I. Knight⁹ & Elena Geangu²  

Sustained attention (SA) is a critical cognitive ability that emerges in infancy and affects various aspects of development. Research on SA typically occurs in lab settings, which may not reflect infants' real-world experiences. Infant wearable technology can collect multimodal data in natural environments, including physiological signals for measuring SA. Here we introduce an automatic sustained attention prediction (ASAP) method that harnesses electrocardiogram (ECG) and accelerometer (Acc) signals. Data from 75 infants (6- to 36-months) were recorded during different activities, with some activities emulating those occurring in the natural environment (i.e., free play). Human coders annotated the ECG data for SA periods validated by fixation data. ASAP was trained on temporal and spectral features from the ECG and Acc signals to detect SA, performing consistently across age groups. To demonstrate ASAP's applicability, we investigated the relationship between SA and perceptual features—saliency and clutter—measured from egocentric free-play videos. Results showed that saliency in infants' and toddlers' views increased during attention periods and decreased with age for attention but not inattention. We observed no differences between ASAP attention detection and human-coded SA periods, demonstrating that ASAP effectively detects SA in infants during free play. Coupled with wearable sensors, ASAP provides unprecedented opportunities for studying infant development in real-world settings.

Keywords Sustained attention, Infant development, Computational model, Electrocardiogram, Wearable sensors, Naturalistic studies, Visual saliency

Sustained attention (SA) is a fundamental cognitive ability that emerges during infancy and has a widespread impact on various developmental domains^{1–3}. As a form of endogenous attention, SA refers to the ability to focus on a particular stimulus or event over an extended period, even with distractors present⁴. SA shapes infant learning^{5–7} and is associated with the emergence in infancy and early childhood of more complex cognitive and social abilities, such as executive function^{5,8}, self-regulation^{4,9–12}, memory^{1,13,14}, language, and social communication^{15,16}. However, most developmental SA research relies on lab-based paradigms that fail to capture the complexity of infants' day-to-day environment^{7,9,17}. Infants' everyday experiences, including the seamless flow of events and social interactions, can vary enormously both within and across cultures^{18–22}. By studying these free-flowing events, we can understand the impact of the stimuli-rich nature of real-life situations on infants' attentional processes, as well as the similarities and differences with the paradigms implemented in the laboratory^{23–25}. It is also possible that phenomena otherwise not observed or studied in the laboratory may occur in the real-world contexts, and systematic observations based on multimodal data may lead to their discovery.

¹Department of Psychological and Cognitive Sciences, Tsinghua University, Beijing 100084, People's Republic of China. ²Department of Psychology, University of York, York YO10 5DD, England. ³Department of Psychology, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, England. ⁴School of Physics, Engineering and Technology, University of York, York YO10 5DD, England. ⁵Bristol Medical School, University of Bristol, Oakfield House, Bristol BS8 2BN, England. ⁶Bioscience Institute, Newcastle University, Newcastle Upon Tyne NE1 7RU, England. ⁷School of Psychology, Newcastle University, Newcastle Upon Tyne NE1 7RU, England. ⁸Department of Psychology, University of Aberdeen, Aberdeen, UK. ⁹Department of Mathematics, University of York, York YO10 5DD, England. ¹⁰Yisi Zhang, A. Priscilla Martinez-Cedillo and Harry T. Mason equally contributed to this work. ✉email: elena.geangu@york.ac.uk

The development of affordable, lightweight and user-friendly wearable technology for infants has facilitated the recording of what infants see and hear in natural settings, and related autonomic nervous system (ANS) responses²⁶. These recordings have great potential for precise measurement and characterisation of SA development as a function of mutual interactions with various factors²⁷. However, a key challenge in adopting ecologically valid approaches is the lack of efficient methods for extracting reliable measures of SA from the available modalities and the vast amount of data²⁸. In this study, we reduce this bottleneck by introducing an innovative automatic attention detection algorithm that harnesses infant electrocardiography (ECG) and accelerometer (Acc) signals recorded with wearable sensors.

Traditionally, looking times have been the predominant measure for visual attention in infancy^{29–31}. Trained researchers can determine whether infants are looking either towards or away from stimuli in various set-ups (e.g., computer-based presentations; social interactions), provided there is an accurate view of the infant's eyes relative to the stimuli of interest³². Due to this approach's simplicity and the relatively low cost of the necessary technology, looking times have been widely adopted in infant development research, enabling the investigation of many important questions. Consequently, it is unsurprising that, in recent years, algorithms have been developed to automate and enhance this approach³².

However, evidence from brain imaging and psychophysiology shows that not all infant looking times reflect active cognitive processing. Specifically, continuous periods of looking estimated by trained researchers can encompass both visual SA and other attentional processes, and inattention^{14,33,34}. In contrast, heart rate (HR) changes (deceleration and acceleration) have been shown to more effectively differentiate between attentional processes, particularly for visual SA^{7,14,35}, and inattention. Orienting towards a stimulus of interest is typically characterized by rapid HR deceleration relative to baseline or pre-stimulus levels. If attention orientation is followed by SA for further processing, HR stabilises at the lower rate for 2–20 s, but can be even longer³⁵. This reduced HR is often also accompanied by decreased body movement³⁶. Attention termination (disengagement from processing the stimulus of interest) is marked by a rapid HR acceleration to approximate pre-attention baseline levels⁵. HR-defined measures of infant SA (HRDSA), but not generic bouts of looking have been associated with neural indicators of active attention and in-depth information processing such as frontal electroencephalogram (EEG) synchronisation in theta oscillations, an effect observed even in infants as young as 3 months^{5,7}. These findings support HRDSA as a robust method for studying the development of SA in infancy.

Building on HRDSA, we developed an automatic SA prediction (ASAP) method to detect attention periods using ECG and Acc signals obtained from wearable sensors commonly used in infant research^{26,37,38}. To train and validate this method, we created a ground-truth dataset consisting of ECG and Acc signals, human-coded SA annotations, and eye-tracking data collected in the laboratory which was set up to emulate home situations (e.g., play mat with toys). ASAP can thus be applied to wearable sensor data collected in natural environments, providing a non-invasive means of studying infant SA. One potential application involves combining ASAP with recordings from lightweight, wireless head-mounted cameras. These cameras, which are well tolerated by infants, can continuously capture the diversity of visual information present in their egocentric views over extended periods of time. However, on their own, these data may not clearly indicate when visual information is likely to be processed in depth. The integration of HRDSA derived from wearable body sensors synchronised with head-mounted cameras can overcome these limitations, enabling researchers to differentiate between egocentric views during which infants manifest sustained attention vs. inattention.

ASAP can be decomposed into three primary steps, motivated by the abrupt HR deceleration and acceleration that occur during attention orientation and termination, immediately before and after the period of SA. The first step involves change point detection (CPD)^{39,40}, which identifies time points when the statistical properties of a signal undergo significant changes. Infancy and toddlerhood are characterised by frequent spans of SA^{41,42}, suiting a CPD algorithm which can objectively identify multiple change points within HR time series data. In the second step, momentary attention detection is formulated as a binary classification task. The classification model is trained with a curated set of features using a feature selection process. In the final step, the segmentation of SA periods is refined to preserve their temporal structure.

The temporal and spectral features of HR fluctuations provide rich information about the ANS' regulation of homeostasis, physiological arousal, and cognitive states^{43,44}. Precisely, heart rate variability (HRV)—measured through variations in beat-to-beat intervals in the time domain or through low (0.04–0.15 Hz) and high-frequency (0.15–0.4 Hz) HR oscillations in the frequency domain—is associated with sympathetic and parasympathetic nervous activity⁴⁵. Higher HRV is linked to better performance in SA tasks (e.g., continuous performance tests involving executive functions⁴⁶). Low-frequency HR oscillations are related to attentional demands^{47,48}, and an increased low-to-high frequency power ratio is associated with poor attention in children with attention deficit hyperactivity disorder (ADHD)⁴⁹. We characterise HR fluctuations in both time and frequency domains using a wavelet transform to capture dynamic HRV changes, potentially relating to momentary shifts in attentional states. Wavelet analysis has been widely applied to characterise instantaneous features of physiological signals^{50–52}. Given the tight coordination between movement and cardiac output⁵³, we also include Acc signals and investigate their dynamic coupling with HR in time and frequency for additional attention insights.

To demonstrate ASAP's utility in attention research, we apply ASAP in a study examining what drives SA when infants and toddlers actively engage with people and objects. Attention allocation can be influenced by factors such as low-level visual features (e.g., strong edges, bright colours, large movements in the scene), people (e.g., faces and bodies), and objects (e.g., toys)^{29,34–64}. However, most previous studies present images or videos under constrained lab conditions (e.g., sitting on a caregiver's lap with stimuli presented on a monitor), which fail to capture infants' and toddlers' active interactions with their environment, and how this may change throughout development. Here we focus on visual saliency⁶⁵ and visual clutter⁶⁶ extracted from observers' egocentric views during free play. Additionally, we examine how these measures differ across fixated regions

during attention and inattention periods, as determined by the ASAP method or human coders, and assess whether they vary with age.

Methods

Overview of approach

Our approach to developing an automatic SA-detection method hinges on two pivotal ideas: first, creating a lab-based ground-truth dataset with human-coded SA; second, training a model using this dataset to automatically detect SA based on ECG and Acc signals.

The multimodal dataset comprises four synchronised sensor signals: ECG, Acc, the scene videos, and fixation-label data. ECG and Acc signals are simultaneously captured from a single wearable device attached to the infant's chest. Fixation data are obtained from a head-mounted eye tracker that records infants' and toddlers' gaze points (Fig. 1), which determine the objects fixated upon during recording. HR is derived from the ECG signal^{26,67}. We annotated attention periods synchronised with the HR and Acc signals by considering both the distinctive HR waveforms during periods of SA and human inspection of aligned fixation data (Fig. 1, "Human Coding").

The model development uses HR and Acc signals alongside human-coded attention (Fig. 1, "Attention Prediction") and it involves three key stages: (1) coarse segmentation of potential periods of SA based on HRDSA facilitated by CPD; (2) classification employing HR and Acc features; and (3) fragment refinement to reconstruct the temporal structure. The model's innovation lies in several aspects: first, the integration of a CPD

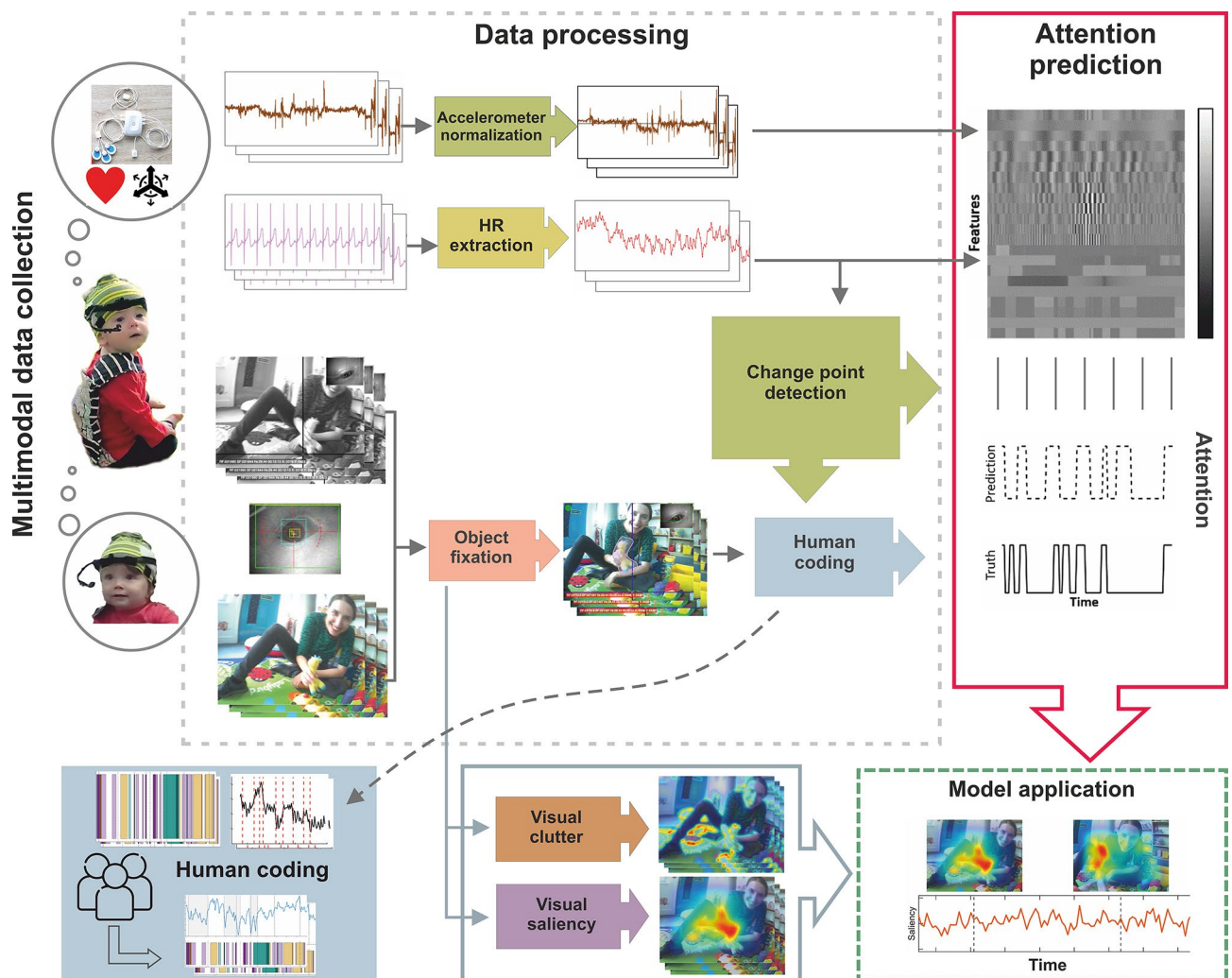


Fig. 1. A pipeline schematic for data collection to train the SA prediction model. Two synchronised wearable devices record the data: the ECG/Acc body sensors (top cloud) and the head-mounted eye tracker (bottom cloud). During data processing, HR is extracted and then undergoes change points detection (CPD) to facilitate human coding of attention, validated by object fixation (objects coded by colours). During attention prediction, HR- and Acc-derived time series and change point segmentation form the feature matrix for training a machine learning model to predict attention periods. During model application, visual clutter and saliency signals are extracted from video frames to evaluate the model's effectiveness.

algorithm to objectively identify abrupt changes in HR, adapting previous lab-based procedures³⁵ to naturalistic settings; second, the exploitation of temporal and spectral properties of HR and Acc signals and their interaction to delineate attentional states; and finally, preserving the temporal structure to retain the associated statistical characteristics throughout the attention detection process. We then demonstrate the model's application in studying visual SA development by leveraging the egocentric video and fixation data collected with the HR and Acc signals (Fig. 1, “Model Application”).

The dataset

Participants

In line with previous literature⁶⁸, a sample size of $N = 75$, 6- to 36-month-old infants and toddlers were included in the final analyses (Table 1). A further 23 participants responded to the invitation to participate in the study but were excluded from the final analysis due to either refusal to wear at least one of the devices ($N = 15$) or technical errors ($N = 8$). Participants were recruited from the urban area of York, in England. Caregivers provided written informed consent before the experimental procedure began, and families received £10 and a book. For those participants depicted in figures, caregivers provided consent for publication of identifying information/images in an online open-access publication. The research presented in this empirical report received ethical approval from the Department of Psychology Ethics Committee, University of York (Identification Number – 119). The experimental procedures were conducted in adherence to the principles of the 1964 Declaration of Helsinki 8 and its later amendments.

Data acquisition and processing procedure

The experiment took place in a laboratory room with controlled lighting containing toys and objects, and a separate larger play area resembling a typical playroom with age-appropriate toys and books (Fig. S1). There were three tasks: check-this-out game, spin-the-pots task, and free play. Participants wore a head-mounted eye tracker to record scene video and gaze points, and body sensors to record ECG and Acc. Further details are provided in the Supplementary Information (1. Task description).

Head-mounted eye tracker data recording and processing

Eye movements were recorded using a head-mounted eye tracker (Positive Science, New York, USA), which tracked the right eye at 30 Hz. Scene recordings were captured at 30 fps with 640×480 -pixel resolution and a wide lens ($W\ 81.88^\circ \times H\ 67.78^\circ \times D\ 95.30^\circ$). The Yarbus software (version 2.4.3, Positive Science) was used to map participants' gaze points onto the scene video and calibrate the eye tracker offline, accounting for variations in eye morphology and the spatial location of fixated objects^{41,69} (further details in the Supplementary Information—2. Head-mounted eye-tracker calibration protocol). We calculated fixations from gaze points and labelled them using the GazeTag software (version 1, Positive Science). We only include fixations with a duration > 100 ms. The labels represent 87 different items including toys, body parts, and other objects in the room. Approximately 10% of frames were unsuitable for labelling due to abrupt movements, participants removing the eye tracker, or technical errors.

Cardiac activity and body movement recording

ECG and Acc signals were recorded concurrently at 500 Hz using the Biosignalsplux device (PLUX Biosignals, Lisbon, Portugal). The ECG sensor's three-electrode montage (including ground) was attached to the left side of the chest, and the Acc sensor was placed at roughly the same location (Fig. 2). The device also includes a light sensor, enabling synchronisation of ECG/Acc data with the eye-tracking data. To ensure quick sensor placement, the Biosignalsplux hub and sensors were embedded in a custom-made shirt (Fig. 2c). Previous infant studies have also considered the movement of the head and/or limbs using various methods (e.g., behavioural videography, desk-mounted eye-tracking, wireless sensors)^{53,70,71}. As the accelerometer available to us was tethered, we decided to only include the device placed on the chest. An additional head accelerometer would restrict head-movement due to additional cables and increase both the preparation time and the number of visible pieces of equipment, which could negatively impact participant compliance.

Sensor signal preprocessing and HR extraction

The triaxial accelerometer (Acc) data was centred by subtracting 2^{15} and then normalised by dividing by 5000. The magnitude of acceleration was calculated as the L2-norm of the triaxial acceleration.

ECG processing and R-peak detection was handled using Python^{26,67}, adapted from Neurokit2 code⁷². All R-peaks were then visually inspected; mislabelled peaks were manually corrected. The HR was calculated using the time interval between consecutive R-peaks (Δt_{peaks}): $HR\ (bpm) = 60/\Delta t_{\text{peaks}}$. The detailed procedure for

Age	N (Total = 75)	M age (months)	SD age (months)	Mean duration of recording (minutes)	SD duration (minutes)
6-months	15 (6 females)	6.43	0.48	25.09	6.81
9-months	15 (8 females)	9.17	0.47	26.27	7.58
12-months	15 (7 females)	12.47	0.99	25.78	4.62
24-months	16 (8 females)	25.50	1.98	29.99	7.55
36-months	14 (8 females)	36.16	2.86	32.51	7.22

Table 1. The cohort breakdown of the participants in the Dataset.



Fig. 2. The wearable body sensor. (a) Example of a 9-months-old infant wearing the head-mounted eye tracker with the body sensors. (b) The Biosignals Plux wearable sensors: 1—the data recording hub; 2—the light sensor; 3—the acceleration sensor; 4—the ECG sensor with the Ambu blue electrodes (Ambu, Copenhagen, Denmark) attached; 5—the custom-made shirt. (c) The back view of the shirt showing how the sensors were embedded: 1—eyelet, which allows the ECG and Acc sensors to be brought from the back to the front and positioned on the left side of the chest; 2—the back pocket, which holds the data recording hub; 3—the shirt closes at the back via hook and loop. (d) Front view of the shirt with the sensors embedded, illustrating the location of the ECG sensor.

noise correction has been described elsewhere⁶⁷. The parameters used in this study are in the Supplementary Information (3. *ECG signal preprocessing and HR extraction*).

Bluetooth connectivity issues caused ECG/Acc data loss six times in total across all 75 participants, with a mean data loss duration of 159.1 s ($SD=106.5$ s; range = 28.9–279.5 s). No HR was calculated during these periods.

Human coding of sustained attention

We adapted the lab-based criteria for HRDSA for data acquired under our more naturalistic conditions. In particular, we defined SA as: (Criterion 1) occurring during periods with a deceleration in participants' HR followed later in the period by an HR acceleration^{35,73–77}; and (Criterion 2) when participants fixate on and/or engage with a small number of objects during this time⁷⁸. Detecting SA in naturalistic settings, such as free play, presents challenges compared to controlled lab conditions. Infants and toddlers in the lab have restricted movements and their HR is measured during a baseline period before an engaging stimulus is presented on a screen. However, free play lacks defined event structures and infants and toddlers can actively fixate and interact with objects, making it difficult to determine baseline HR measurements.

To address these challenges, we first developed a method to identify candidate periods corresponding to Criterion 1. We used a CPD algorithm to automatically detect abrupt decreases (SA onsets) and increases (SA terminations) in HR without the need for a pre-determined baseline (see *The Automatic Sustained Attention Prediction (ASAP) Model* section). This provides an objective approach to detect abrupt changes by adaptively responding to the data, minimising the need for subjective, user-defined criteria, and holding the potential to accommodate variations attributable to individual differences, fluctuations in HR baseline, and the developmental change of HR. Second, we use eye-tracking data to determine whether these periods were associated with infants fixating on or following objects within the scene corresponding to Criterion 2.

Human coding protocol

Three experimenters (two authors) independently coded SA periods. Figure 3 illustrates the custom MATLAB graphical user interface for the protocol. First, detected change points in the HR time series were marked to create segments (Fig. 3a, vertical dashed lines). The change in mean HR of a segment relative to the preceding segment was then calculated.

The change in mean HR on consecutive segments was next screened for putative SA periods as follows. The SA onset was set to the change point when there was a decrease of at least 5 bpm in the mean HR relative to the preceding segment. If the change was between 3 and 5 bpm, the coder assessed the peak-to-peak drop around

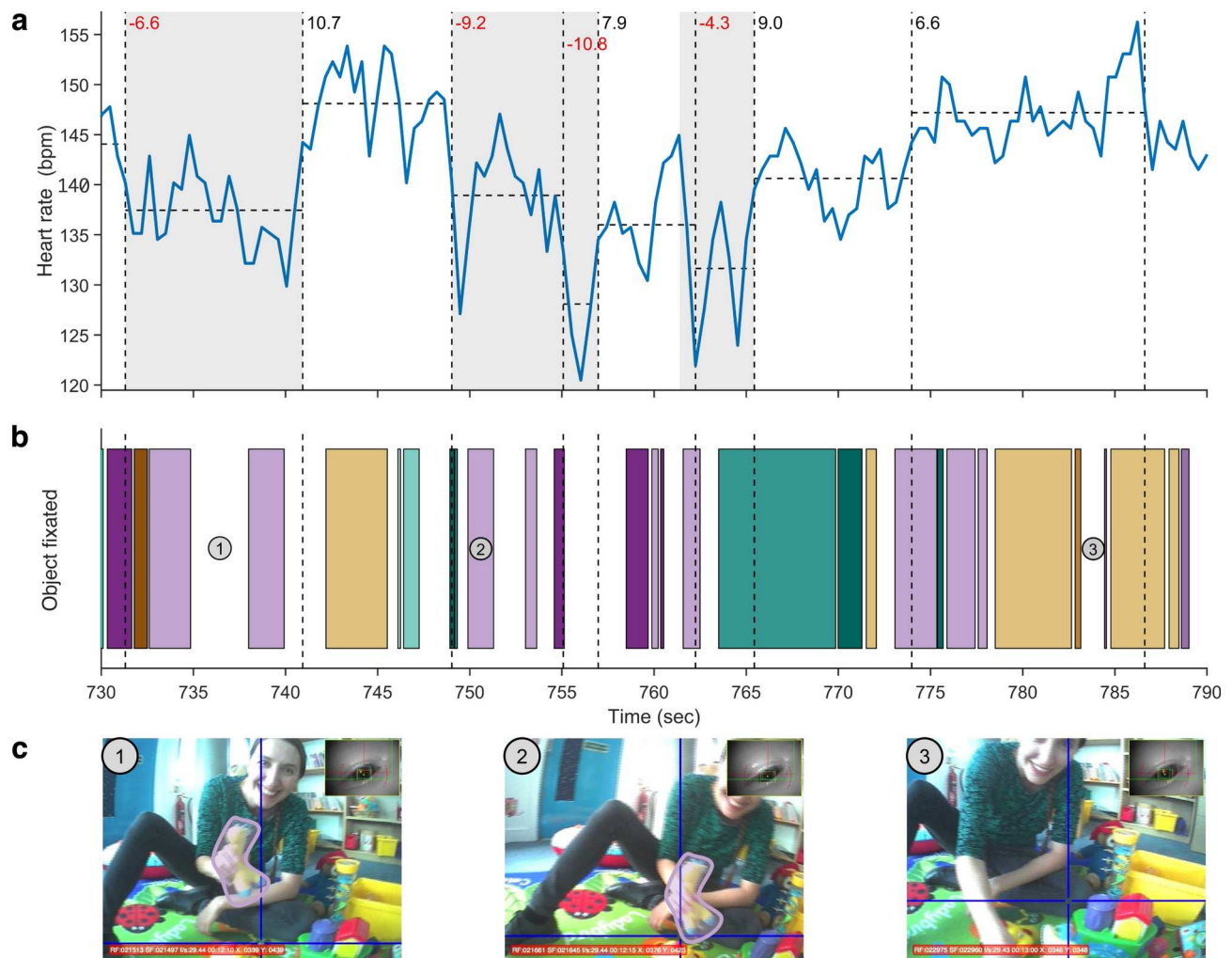


Fig. 3. Example of human-coded sustained attention for a 6-month-old participant for a 1-min time window. **(a)** The solid blue line indicates the HR time series (in bpm). The vertical dashed black lines indicate change points which divide the HR time series into segments. The horizontal dashed black lines show the mean HR for each segment. The change in mean HR relative to the preceding segment is displayed in the top left of each segment (red indicates deceleration; black indicates acceleration). The grey regions indicate sustained attention periods. In the third grey region, the onset time was shifted to the first HR peak before the change point as the change in HR was -4.3 bpm, with the change between the HR peak immediately before and after the change point > 5 bpm. **(b)** The fixated objects from the eye-tracking data. Each unique colour bar represents a unique object (57 total objects), e.g., purple bars represent periods of fixation on a plush giraffe. **(c)** Three representative frames from the infant's egocentric view were extracted at each point in time corresponding to the numbered grey circles from Panel B. The fixation (blue crosshair) and participant-eye view from the eye tracker are superimposed onto each frame. The plush giraffe outlined and shaded in purple corresponds to the purple bars in Panel B (periods when the infant fixates on the giraffe).

the change point. If the drop exceeded 5 bpm, the onset was set to the time of the peak immediately before the change point (e.g., Fig. 3a, third grey region). For consecutive drops meeting these criteria, the earliest drop determined the SA onset (e.g., Fig. 3a, second grey region). The termination of SA was set at the change point with an HR increase of at least 5 bpm. For 3–5 bpm increases, termination was set to the nearest HR peak after the change point if the peak-to-peak rise was 5 bpm or more. SA periods shorter than 2 s were excluded⁷⁸.

Lastly, coders assessed whether participants were looking at or following a small number of objects during the identified putative SA periods. This was done using the fixation-label time series (Fig. 3b) and inspecting the screen video overlaid with the fixation crosshair (Fig. 3c). The experimental setup contained several toys or objects of potential interest (maximum 25), similar to a typical home^{79,80}, and participants were free to move about the space. In most lab-based studies investigating HRDSA, complex and dynamic stimuli, which comprise multiple social and non-social elements and changes in scenes, were presented on computer screens^{17,81–84}. In these studies, periods of HR deceleration indicative of SA have been associated with both brief and extended 'looks' towards the screen, each 'look' encompassing fixations towards different objects, as well as different scenes and events. Other lab-based studies presented infants with 1 to 6 toys, often handed to them one at a time^{42,78,85}.

Considering the more naturalistic conditions in our study, we set the threshold to a maximum of 5 objects to be fixated or followed in order for a period of HR deceleration to be considered SA.

To assess inter-rater reliability, we randomly selected 15 participants from the five age groups, with two of the three coders independently coding each participant. Reliability was determined by counting overlapping attention and inattention periods between the two coders, allowing for any degree of overlap. The coders showed substantial agreement (Cohen's κ : $M = 0.764$; $SD = 0.116$; range: 0.496 to 0.895).

Saliency and clutter extraction

We calculated saliency and clutter from the acquired scene video frames and fixation data during the naturalistic free-play periods. These measures can attract infant and adult participants' attention^{54,61,64,86–88}. The fixation distribution per age group is presented in the Supplementary Information (Fig. S2, Table S1).

Figure 4 illustrates how we generated a saliency time series (see also Fig. S3). Based on the duration of SA periods in our data, we first segmented the frames into consecutive 5-s time windows to allow for the temporal integration of fixated visual information. Second, we created a 50-pixel radius circular mask^{5,89,90} centred on each fixation and accumulated these masks within each window to create a binary fixation mask. On average, approximately six fixations contributed to each binary fixation mask, covering approximately 7% of the frame area (Table S2). Third, we extracted saliency maps from the frames, applied the binary fixation mask to each map within that window, and averaged the saliency values of pixels inside the mask to create the time series. During time windows with no fixations, the saliency/clutter values were removed from further analyses. We used the Graph-Based Visual Saliency (GBVS) algorithm to compute saliency⁶⁵ and the Feature Congestion measure to compute clutter⁶⁶. Full details of this procedure are provided in the Supplementary Information (5. Saliency and clutter extraction).

The automatic sustained attention prediction (ASAP) model

ASAP detected HRDSA by identifying abrupt changes in HR and leveraging information from HR and Acc signals associated with SA. Using the processed HR and Acc signals, we conducted feature extraction and feature selection to establish a classifier for attention prediction. As shown in Fig. 5, the ASAP procedure consists of three steps (colour-coded in Fig. 5): *change point segmentation*, *point-wise classification*, and *segmentation refinement*.

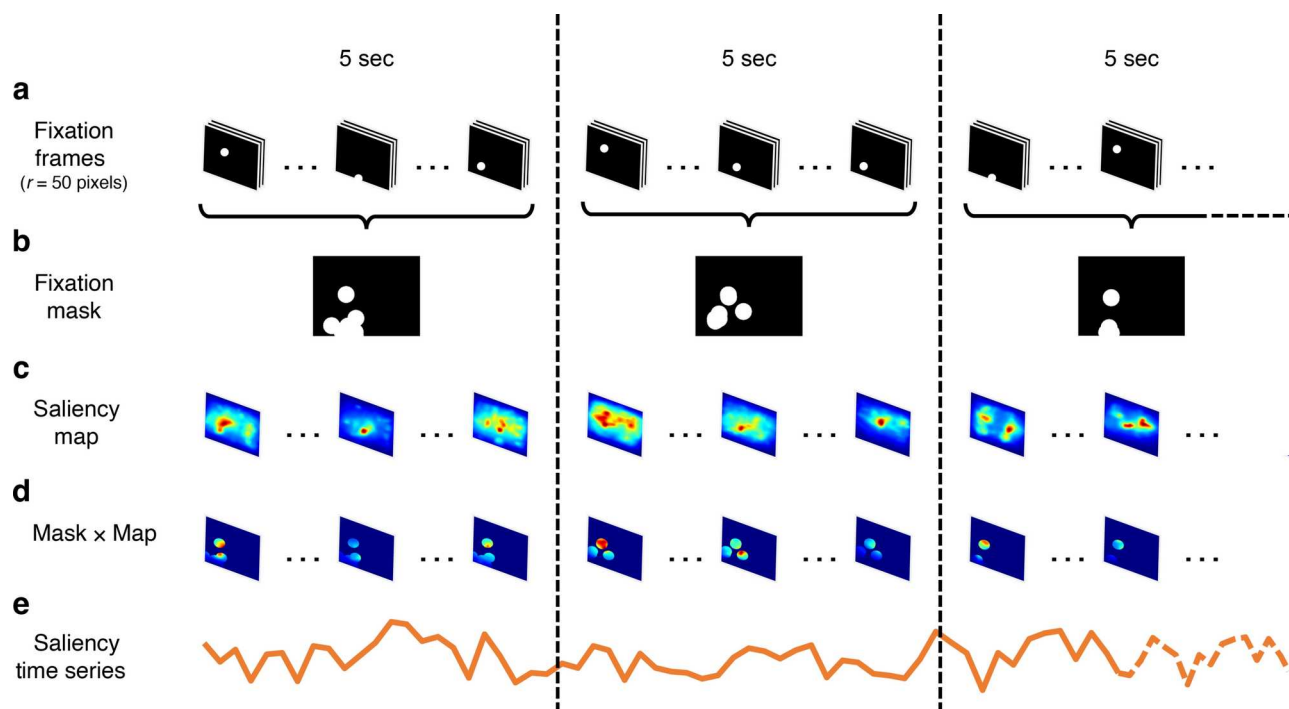


Fig. 4. Example of a saliency time series during free play. (a) The fixation time series is divided into consecutive 5-s time windows (150 frames at 30 Hz). A spatial window was created around each fixation (radius, $r = 50$ pixels; white circles). (b) All fixations within each 5-s time window are accumulated to form a binary fixation mask (black pixels = 0, white pixels = 1). (c) The saliency maps within each window are extracted (15 frames at 3 Hz). (d) Each saliency map within a 5-s time window is multiplied by the corresponding binary fixation mask for that window. (e) The saliency time series is created by averaging across all pixels within the respective binary fixation mask at each time point. The same procedure is used for the clutter time series. For a detailed illustration of the binary fixation mask applied to a saliency map, see Figure S3.

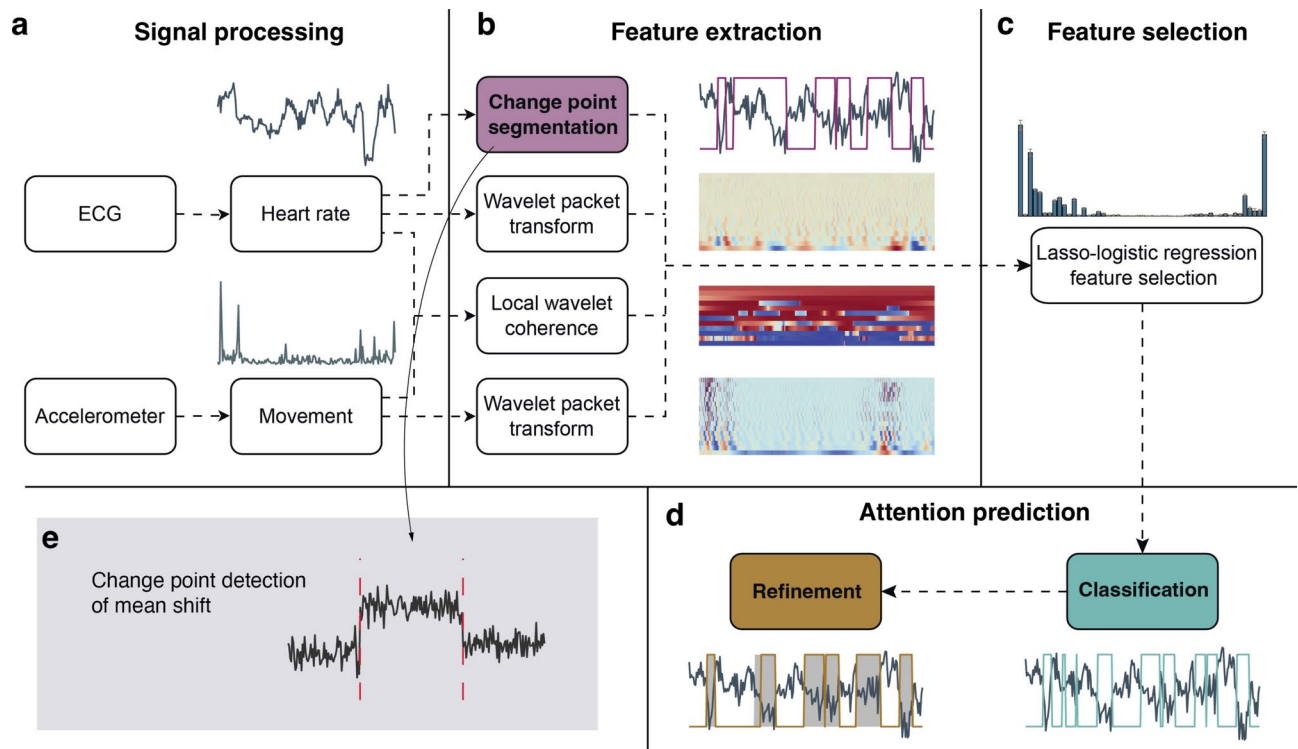


Fig. 5. ASAP procedure overview. **(a)** Signal processing to extract heart rate (HR) and movement (acceleration magnitude). **(b)** Feature extraction to produce change point segments, wavelet packet transforms of HR and Acc, and local wavelet coherence between HR and Acc. **(c)** Feature selection using Lasso regularised logistic regression. **(d)** Attention prediction using logistic regression with selected features and further refinement to reconstruct the temporal structure. Coloured boxes in **(b)**–**(d)** indicate the three stages of attention prediction. **(e)** An example of change point detection of mean shift.

Step 1: Change point segmentation

Change point detection (CPD) is a statistical method that detects changes in properties (e.g., mean, variance, slope) in a time series. Here, the change point segmentation step (Fig. 3) identified boundaries of putative SA periods indicated by steep decreases and increases in the mean of HR (Fig. 5e). Given that infant free play is characterised by frequent short periods of SA and that the heartbeat defines a fundamental minimum resolution of 2–3 samples per second for observing changes, we adopted a CPD methodology that produces change points best fitting these resolution constraints: wild binary segmentation (WBS2) with a model selection criterion of the steepest drop to low levels (SDLL)⁴⁰ (for the selection motivation, see Supplementary Information – 6. *Change Point Detection*). The method is available from CRAN via the R package ‘breakfast’ (version 2.2), using ‘wbs2’ and ‘sdll’ options.

Applied to our dataset, this approach had an average rate of CPD of one change point per 8 to 9 heartbeats. Change points were classified as descending and ascending by calculating the local HR slope. We then applied the same rules as for the *human coding protocol* to screen and refine the boundaries of SA. Step 1 resulted in a binary time series indicating putative attention and inattention segments.

Step 2: Point-wise classification

Feature extraction. Time points were mainly classified on the temporal and spectral information extracted from HR and Acc and the putative SA segments identified in Step 1. Fifty-one predictor features were extracted (see Table S3 and Fig. 6 for full list). (1) The HR and Acc magnitude. (2) Time–frequency decomposition using the wavelet packet transform was used to analyse HR and Acc over different time scales (MATLAB 2022a routine ‘modwpt’, Daubechies wavelet with two vanishing moments), yielding sixteen features (frequency bands) for each measure (WPT-HR and WPT-Acc). (3) The time-evolving spectral cross-dependence between HR and Acc was estimated using multivariate locally stationary wavelet processes, with the localised coherence Fisher’s-z transformed (R package ‘mvLSW’ version 1.2.5)⁹¹, yielding 12 features (LSW-HR-Acc). (4) HRV measures from the recording session were included, specifically the standard deviation of inter-beat intervals (SDRR) – the time deviation between consecutive R-peaks. (5) A binary time series derived from the CPD step was included, indicating potential attention and inattention states (Change point binary). (6) Segment duration (Duration) and time to the previous segment (Latency) yielded two more features. (7) Lastly, age in months was included.

These variables formed the predictors, X , in the model $Y=f(X)$, where Y represents a binary time series of attention and inattention from human coders. We initially included the above features as predictors, before

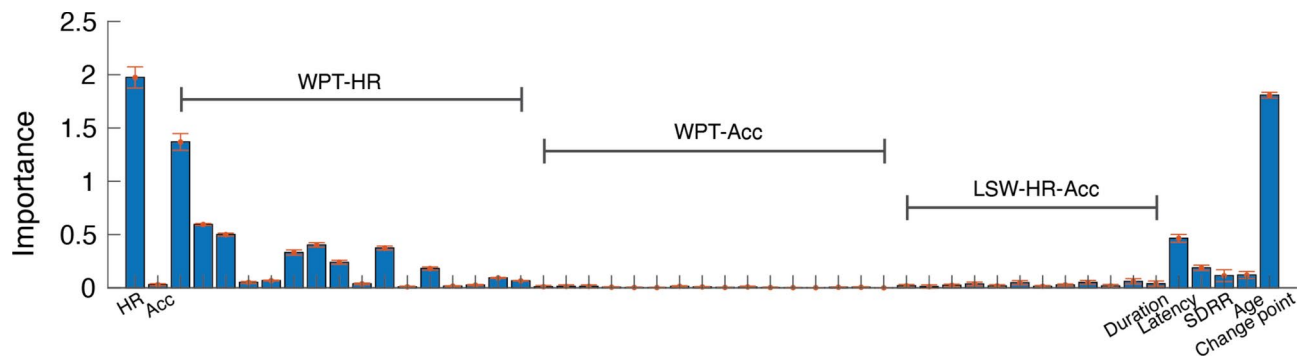


Fig. 6. Feature importance by Lasso regularised logistic regression. Importance is measured using the absolute values of the standardised logistic regression coefficients, averaged across the fivefold cross-validation. Error bars are standard deviations across the fivefold cross-validation. Selected features had mean Importance > 0.03 (threshold determined by cross-validation). WPT-HR: wavelet packet transform of HR; WPT-Acc: wavelet packet transform of Acc; LSW-HR-Acc: local stationary wavelet estimated coherence between HR and Acc. In these cases, each bar corresponds to a frequency band.

performing feature selection to determine the most informative features in predicting attention. All variables were sampled at 2 Hz.

Feature selection. The 75 participant sessions were split into training and test sets (~9:1), where each partition contained randomly selected sessions with approximately equal age distribution. The training set consisted of 67 participant sessions (194,360 time points), and the test set consisted of 8 participant sessions (22,367 time points). Our model selection procedure was conducted using the training set, while the test set was kept separate and only used for the final evaluation of model performance. Feature selection was performed using regularised logistic regression, which shrinks a subset of the estimated coefficients to zero by employing an L1-norm penalty (Lasso) for covariates deemed to have non-significant contributions to attention⁹² (see Supplementary Information, Sect. 7, for the mathematical formulation). The tuning parameter (λ) that controls the strength of the L1-norm penalization was determined by cross-validation ($\lambda = 3.4e - 5$). The regularised logistic regression model was trained using the ‘*lassoglm*’ MATLAB routine with 100% L1 penalty.

The Lasso approach preserved 24 out of the 51 feature variables (Fig. 6 and Table S3), including the HR, a subset of wavelet transforms of the HR (WPT-HR in Fig. 6; 13/16 selected; see Table S3 for selected frequency bands), duration, latency, SDRR, change point binary, and age. The Acc alone and its wavelet transform (WPT-Acc) were determined to be unimportant, but five of the twelve HR-Acc coherence bands (LSW-HR-Acc) were included. The logistic regression model was then retrained without regularisation using the training set based on the 24 features. The performance was then evaluated using the test set. Using other classifiers resulted in comparable performance (Fig. S4, Table S4).

Model training. Imbalanced class distribution, where one class significantly outnumbers the other, can bias model training. In our case, attention occupied about 29% of the total time. To address this, we tested whether balancing class distribution could improve model performance using fivefold cross-validation within the training set. Data balancing was achieved using the Synthetic Minority Over-sampling Technique (SMOTE)⁹³. SMOTE improved the model performance based on the F1-score criterion. We showed that the model’s performance remained stable regardless of dataset size, with performance metrics plateauing when trained on more than half of the training set (see Supplementary Information, Sect. 10, for sensitivity test; Fig. S5). To optimise the model accuracy so that it can be deployed for future data collected in natural environments, we provide the final model trained on the entire dataset with SMOTE oversampling applied.

Step 3: Segmentation refinement

Point-wise classification (Step 2) did not preserve the temporal structure of the putative SA segments obtained in Step 1, and could disrupt long SA periods. To maintain the temporal structure, we merged adjacent segments based on the information from Step 1 and Step 2. First, segments identified in Step 2 that were 2.5 s or less apart were merged. Next, if the predicted attention in Step 2 covered more than 90% of a Step 1 SA segment, the entire segment was classified as an SA period, provided that no attention span exceeded 50 s. These criteria were established using fivefold cross-validation within the training set, which optimised on producing an attention duration distribution similar to the human-annotated distribution. The Kolmogorov–Smirnov (KS) test was applied to compare the model distribution to the human distribution.

Model evaluation

The performance of the ASAP model was evaluated on the test set (8 sessions from all age groups of 6-, 9-, 12-, 24-, and 36- months). First, we assessed the point-wise performance of the method in comparison to human coding for different metrics: accuracy $\frac{\text{\#ofCorrectPredictions}}{\text{Total\#ofPredictions}}$, precision $\frac{\text{TruePositives}}{\text{TruePositives} + \text{FalsePositives}}$, recall $\frac{\text{TruePositives}}{\text{TruePositives} + \text{FalseNegatives}}$, F1-score $\frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$, and inter-rater reliability (Cohen’s κ) at each step. Second, the KS test was used to assess the similarity of the predicted duration distribution compared to the

human-coded data. Finally, we examined whether our prediction method could preserve the natural statistics associated with attention as identified by human coding (see Model application: Cross-validation of the ASAP model with egocentric visual information section). The criterion of statistical significance was set to 0.05.

Results

Model performance

The ASAP procedure progressively approximated human-coded attention periods through the three steps (Fig. 7a, b). The change point segmentation phase (Step 1) attained a point-wise accuracy of $80 \pm 5\%$ (mean \pm SD), precision of $60 \pm 10\%$, recall of $92 \pm 6\%$, F1-score of 0.72 ± 0.07 and Cohen's κ of 0.76 ± 0.07 . An all-negative prediction, based on the assumption that participants were not paying attention for the majority of the time, resulted in a significantly lower accuracy of 71%. Conversely, an all-positive prediction led to a precision of 29%, consequently producing a null F1-score of 0.45. The failure to achieve 100% recall can be attributed to adjustments made during the human coding process incorporating the video and fixation data. Additionally, the duration of attention segments exhibited a broader distribution than the human-coded distribution ($p = 0.0056$, KS test, Fig. 7c).

The classification step (Step 2) improved the prediction accuracy to $84 \pm 4\%$ ($p = 0.0060$, paired t -test between Step 1 and 2) by increasing precision to $68 \pm 9\%$ ($p = 1.9 \times 10^{-4}$, paired t -test). This step reduced recall to $85 \pm 7\%$ ($p = 0.0025$, paired t -test), indicating a trade-off between precision and recall. Cohen's κ also dropped to 0.66 ± 0.08 ($p = 5.0 \times 10^{-4}$, paired t -test) due to the reduction in recall. Step 2 fragmented long attention periods, resulting in a duration distribution skewed towards shorter durations ($p = 5.7 \times 10^{-5}$, KS test, Fig. 7c). Overall, Step 2 improved the accuracy of point-wise predictions but did not maintain the temporal structure.

The refinement phase (Step 3) merged short segments to reconstruct long attention spans. The point-wise performance remained largely unchanged, with accuracy of $84 \pm 4\%$, precision of $67 \pm 10\%$, recall of $86 \pm 6\%$, F1-score of 0.75 ± 0.06 , and Cohen's κ of 0.74 ± 0.07 . Importantly, this phase restored the duration distribution to be statistically comparable to the human-coded distribution ($p = 0.42$, KS test, Fig. 7c). Therefore, Step 3 was crucial for recovering intact attention periods. Additionally, we did not observe a significant effect of age on any of the metrics at any step (p s > 0.2 , linear regression).

Model application: Cross-validation of the ASAP model with egocentric visual information

An infant's attentional state is linked to their sensory experiences within natural environments. We estimated the changes to visual saliency and clutter of fixated regions over time (see Methods—Saliency and clutter extraction section and Fig. 4) and compared these measurements between ASAP and human-coded attention and inattention periods, while also examining potential developmental changes. Only data from free play sessions with at least

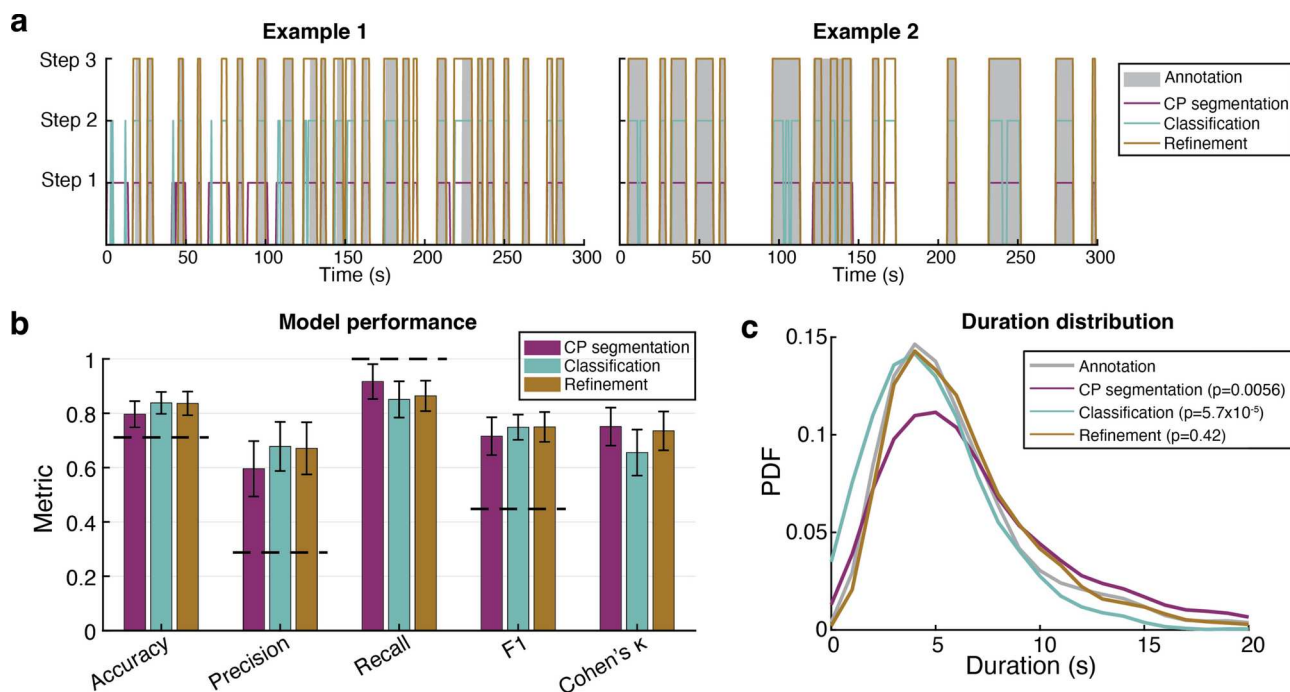


Fig. 7. ASAP performance. (a) Two examples of attention prediction through 3 steps. (b) A summary of performance metrics. Colour codes are identical to (a) and Fig. 5. Dashed lines indicate null estimates. The null is obtained for accuracy when guessing all negative; for precision, recall, and F1, the null is calculated with all-positive guesses. Error bars are standard deviations. (c) Distributions of segment duration are obtained from each step. P -values revealing the distribution similarity are based on KS test. CP: change point. PDF: probability density function.

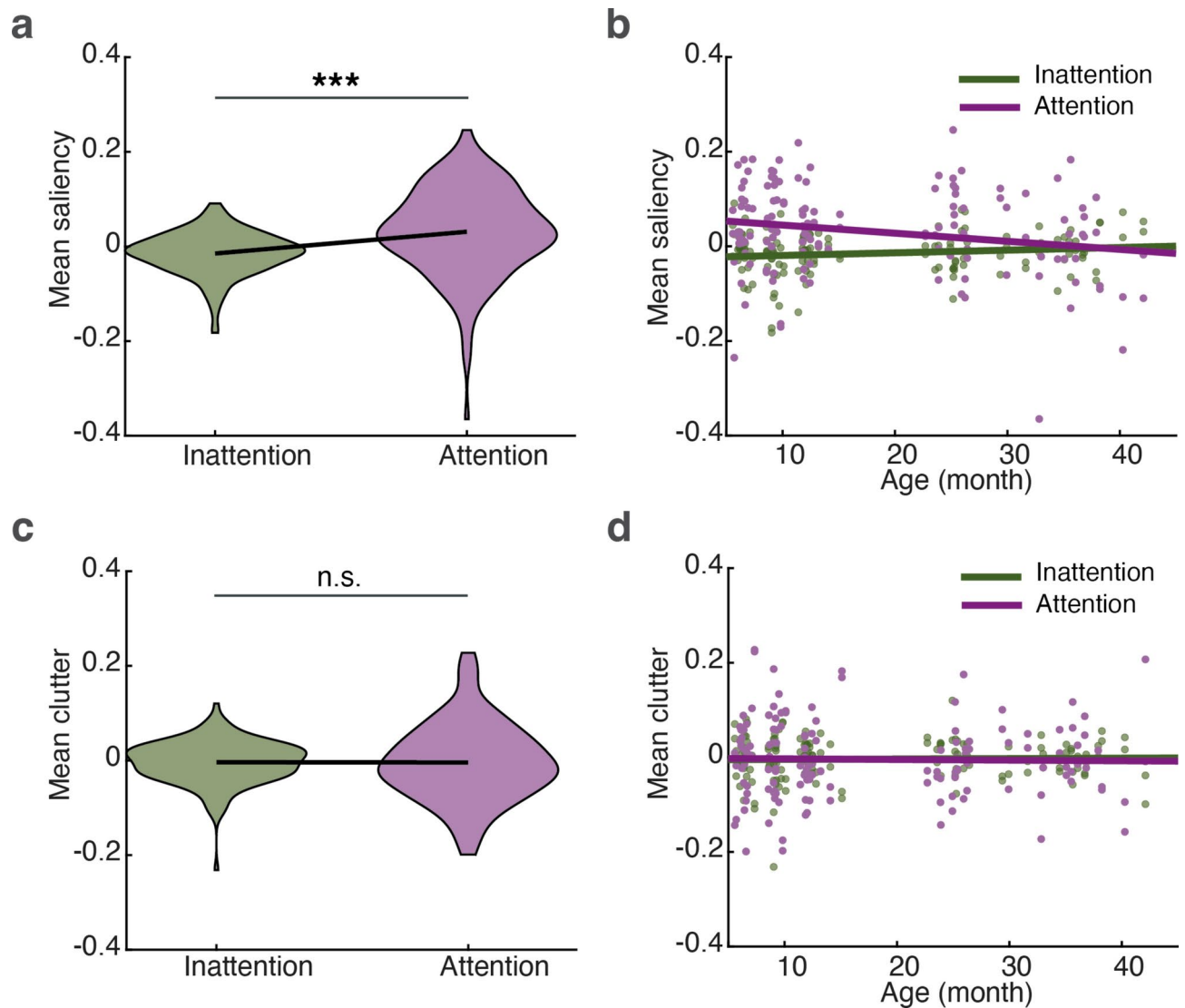


Fig. 8. Effects of attention on visual features. (a) Violin plot of the attentional effect on visual saliency. (b) Interaction effect on visual saliency between attentional states and age, with lines fitted by linear regression. (c) and (d) respectively show the equivalent graphs for clutter.

one minute of cumulative recorded fixation ($N=72$; 966 ± 368 s) were included in this analysis. The number of fixated regions during attention (0.048 ± 0.014 fixations per frame (mean \pm SD)) and inattention (0.047 ± 0.014) did not differ significantly ($p=0.35$, paired t -test). We standardised saliency and clutter measures within each session to eliminate systematic variation across participants. The session means of each measure were analysed using a linear model (*fitglm* in MATLAB, 2022a), with three factors: attentional state (attention vs. inattention), data source (human coding vs. ASAP prediction), and age, along with their four two-way interactions and one three-way interaction term (see Supplementary Information, Table S5). A Bonferroni correction was applied to adjust for multiple comparisons across factors, interactions, and responses.

We found that attentional state had a significant effect on mean saliency, with higher saliency of fixated regions during attention than inattention periods (Bonferroni-corrected $p=9.2e-7$; Fig. 8a). There was also a significant interaction between attentional state and age for saliency (Bonferroni-corrected $p=0.0011$; Fig. 8b): the saliency of fixated regions decreased with age during attention periods but remained relatively constant during inattention periods. No significant effects or interactions involving the data source were found (complete statistical results in Table S5). Additionally, no attentional or age-related effects on the visual clutter of fixated regions were observed (all $ps > 0.05$; Fig. 8c, d; Table S5).

Discussion

There has been increasing interest in employing naturalistic approaches to study human development^{18,94–97}. Sustained attention (SA), a cognitive ability that emerges during infancy, is particularly important due to its widespread implications across various developmental domains^{1,3,5}. Despite its significance, most current SA

developmental research relies on lab-based paradigms, with limited understanding of whether these findings extrapolate to the natural environment, or how the mutual interactions between the developing infant and their everyday environment contribute to SA emergence. In this study, we propose an innovative algorithm—ASAP—that harnesses signals recorded unobtrusively from wearable technology (e.g., the EgoActive platform²⁶) to detect infant SA manifested spontaneously.

ASAP is the first model to utilize ECG and Acc signals to classify attentional states with high temporal resolution in preverbal infants and toddlers, making it particularly well-suited for settings outside the laboratory, where the environment changes continuously and infants are free to move around. Previous efforts for using machine learning methods based on physiological signals (e.g., ECG, EEG) have exclusively focused on predicting attentional states in adults^{98–102}. These studies have predominantly used well-controlled lab-based paradigms, with the ground truth established based on adult participants' self-reported attentional states or the nature of the stimuli/task, usually over longer timescales (several minutes). Other applications of ECG signals in machine learning have aimed to diagnose neurocognitive disorders such as ADHD¹⁰³ or autism¹⁰⁴. Similar to our approach, these studies employed feature extraction, including spectro-temporal decomposition of physiological signals, to train classification models. However, unlike our study, these models were not trained to predict attentional states or specific cognitive states, but rather to predict specific diagnoses. Furthermore, an important innovation of our approach is the integration of change point detection, which enables the robust and efficient classification of densely sampled attention and inattention periods. This innovation provides a powerful tool to capture and understand infants' attentional state in response to a dynamically changing environment.

The key HR features contributing to ASAP's successful SA detection are the HR deceleration during periods of SA relative to those of non-attention³⁵, and the HR fluctuations in both time and frequency domains^{45–48,50–53,105}. The feature importance analysis revealed that HR fluctuations in the 0.03–0.15 Hz and 0.3–0.6 Hz frequency bands contributed significantly to predicting SA (Fig. 6 and Table S3), corresponding to the two distinct peaks previously observed in the infant HR power spectrum¹⁰⁶. A reduction in body movement has also been previously used as a criterion for determining attention^{107,108}, and associations have been shown between changes in HR and body movement linked to measures of attention^{70,71,109,110}. While previous studies have considered the movement of the head or limbs^{70,71}, the feature selection in this study did not identify the torso Acc magnitude as playing a significant role in predicting SA. Our choice for relying on the torso movement was motivated by both empirical and pragmatic reasons. Empirically, comprehensive studies with non-human primates show that when considered alongside movement of the head and limbs, torso movement is most strongly coupled with the HR⁵³. Pragmatically, in many of the wearable sensors typically used for naturalistic recordings, accelerometers are placed on the torso²⁶. Although Acc magnitude did not predict SA in this study, its coupling with cardiac activity did, especially at very low frequencies below 0.1 Hz (Table S3). This frequency range aligns with previously reported spontaneous brain fluctuations that are associated with arousal and its coordination with movement in humans and non-human primates^{53,111,112}. Considering that some of the key features in defining periods of SA are represented by the deceleration and acceleration of the HR, as well as HR fluctuations in the frequency domain, our findings suggest that these features are less likely to be the direct result of the changes in torso movement.

We also investigated whether the variability in performance across sessions was driven by intrinsic dataset characteristics or biases inherent to the model. By analysing the correlation between human coder agreement and the agreement between human coders and the ASAP model, we found that intrinsic factors influenced performance for both (Supplementary Information, Fig. S6a, b). Further analysis indicated that SDRR (an HRV measure) was a key intrinsic feature that contributed to SA classification performance (Fig. S6c). Previous studies have shown that respiratory sinus arrhythmia (RSA), a separate HRV measure within the respiratory frequency range, predicts the extent of HR deceleration entering the attentional phase⁷⁶. Individuals with higher RSA levels (also higher SDRR) are more resistant to distraction^{46,76}. Therefore, higher HR variation may lead to more distinct HRDSA, potentially reducing the confusion in the attention classification task. This correlation supported the inclusion of SDRR as a feature in the attention prediction model, and our feature selection analysis further validated its relevance. Overall, these supplementary analyses demonstrate that the ASAP model did not have any inherent bias leading to systematic errors, and that errors in model classification align with discrepancies between human coders.

We applied ASAP-labelled data to study visual attention development by leveraging the egocentric video and fixation data and generating the perceptual features of visual saliency and clutter. Three key findings emerged. First, and of critical importance, the results were similar for the periods of attention determined by human coders and those detected by ASAP (Table S5). This demonstrates that implementing ASAP can be a less costly (time and human resources) approach, making it ideal for studies involving large data recorded in the natural environment or in the lab. Second, similar to some previous studies¹¹³, during SA periods, infants fixated on areas with higher perceptual salience than during periods of inattention. Older infants were observed to fixate on less perceptually salient regions during sustained attention than younger infants. This could reflect the older infants' increased reliance on more high-level properties of the scenes, although further consideration of other factors (e.g., local meaning) that tend to covary with salience would be required to more definitively conclude this effect⁶¹. Third, unlike some of the previous studies¹¹⁴, the areas fixated during attention and inattention did not seem to differ in terms of visual clutter, irrespective of infants' age. In part, this could be due to the fact that here we considered the amount of fixated feature congestion, which may be less relevant for allocating and sustaining attention compared to how cluttered the entire scene is¹¹⁴. Our results help to replicate and extend findings from lab-based studies under constrained conditions to more naturalistic scenarios. Importantly, for the first time, the relation with visual saliency is reported for HRDSA, supporting the interpretation that visual saliency not only influences what infants are likely to orient towards but also what is likely to be processed in depth^{73,115}.

It is interesting to note that infants' and toddlers' capacity to control the environment can reduce visual clutter in their egocentric view. Older toddlers are more likely to be able to control their visual clutter, whether by interacting with objects directly or by moving their body, head, or eyes. These physical interactions with their environment can also change the visual saliency from their egocentric view. By comparison, caregivers are more likely to move objects in younger infants' environments as infants may have less motor control than older toddlers. We did not find a main effect of age on the overall levels of visual saliency or feature congestion ($ps > 0.14$, see Supplementary Table S5). These null findings may relate to individual differences in motor development, for instance; and these differences may possibly be larger than age differences^{116–118}. Thus, considering both age and individual differences can be an exciting avenue of future research.

Limitations and future research

ASAP is based on HR-defined visual attention. Our model yields a weaker precision (~67%) relative to the other metrics, which may be partly due to other events inducing similar HR changes. Attention in other sensory modalities (e.g., audition) can also be accompanied by HR deceleration¹¹⁹, resulting in an HR waveform similar to that observed during visual attention. Additionally, vocalisations have been associated with an HR dynamic that mirrors the one observed during visual SA^{53,105}. Other information can help differentiate these possibilities from visual SA (the ground truth in our study), and future work could extend ASAP to detect sustained auditory attention while also incorporating infants' and toddlers' vocalisation as features¹²⁰, potentially improving its performance.

This study represents the first step in developing the ASAP algorithm based on a methodological lab set-up that includes both structured and unstructured activities, some of which emulate those carried out at home (e.g., free play on a mat with toys). This step is crucial to ensure the reliability of the algorithm. Future research may consider data recorded in home settings, with a methodological set-up adequate for this context (e.g., wireless wearable sensors such as the EgoActive²⁶, a combination of home- and lab-based assessments, as well as assessments of attention based on parental reports¹²¹).

It would also be interesting to explore the possibility of developing new algorithms that complement ASAP to predict other attentional processes. Studies that integrate eye-tracking and ECG to investigate infant attention responses to emotional stimuli have shown strong associations between a lower frequency of saccades away from emotional faces and concomitant HR deceleration as an index of attention-orienting responses¹²². These findings raise the possibility that ECG signals may also have the potential to be used as predictors for attention orientation and eye-movements.

Conclusions

We demonstrated that ASAP achieved solid performance in the challenging task of detecting periods of sustained attention based on HR and Acc signals across the age range and conditions considered here, including natural, free-play behaviours. This is critical as HR and Acc are the predominant signals recorded with unobtrusive wearable sensors for infants and are invariant across platforms. Although future work can expand ASAP to include other sensory modalities and increase its performance metrics, it can currently provide a powerful tool to detect visual sustained attention in natural settings (e.g., home), where infants and young children are free to move. Together with available wearable technology, ASAP provides unprecedented opportunities for understanding how attention develops in the context of infants' and toddlers' daily experiences, and to support explanations of attention development with high ecological validity.

Data availability

All new data and code required to reproduce the main results and figures of the article are available at: <https://github.com/yisizhang/ASAP>. The raw video files cannot be shared in order to protect participants' privacy and confidentiality, in line with data protection legislation.

Received: 9 December 2024; Accepted: 28 March 2025

Published online: 17 April 2025

References

1. Reynolds, G. D. & Romano, A. C. The development of attention systems and working memory in infancy. *Front. Syst. Neurosci.* <https://doi.org/10.3389/fnsys.2016.00015> (2016).
2. Rose, S. A., Feldman, J. F. & Jankowski, J. J. Attention and recognition memory in the 1st year of life: A longitudinal study of preterm and full-term infants. *Dev. Psychol.* **37**, 135–151. <https://doi.org/10.1037/0012-1649.37.1.135> (2001).
3. Ruff, H. A. & Lawson, K. R. Development of sustained, focused attention in young children during free play. *Dev. Psychol.* **26**, 85–93. <https://doi.org/10.1037/0012-1649.26.1.85> (1990).
4. Casey, B. J. & Richards, J. E. Sustained visual attention in young infants measured with an adapted version of the visual preference paradigm. *Child Dev.* <https://doi.org/10.2307/1130666> (1988).
5. Brandes-Aitken, A., Metser, M., Braren, S. H., Vogel, S. C. & Brito, N. H. Neurophysiology of sustained attention in early infancy: Investigating longitudinal relations with recognition memory outcomes. *Infant Behav. Dev.* <https://doi.org/10.1016/j.infbeh.2022.101807> (2023).
6. Frick, J. E. & Richards, J. E. Individual differences in infants' recognition of briefly presented visual stimuli. *Infancy* **2**, 331–352. https://doi.org/10.1207/s15327078in0203_3 (2010).
7. Xie, W., Mallin, B. M. & Richards, J. E. Development of infant sustained attention and its relation to EEG oscillations: An EEG and cortical source analysis study. *Dev. Sci.* **21**, e12562 (2018).
8. Johansson, M., Marciszko, C., Gredebäck, G., Nyström, P. & Bohlin, G. Sustained attention in infancy as a longitudinal predictor of self-regulatory functions. *Infant Behav. Dev.* **41**, 1–11. <https://doi.org/10.1016/j.infbeh.2015.07.001> (2015).

9. Brandes-Aitken, A., Braren, S., Swingler, M., Voegtline, K. & Blair, C. Sustained attention in infancy: A foundation for the development of multiple aspects of self-regulation for children in poverty. *J. Exp. Child Psychol.* **184**, 192–209. <https://doi.org/10.1016/j.jecp.2019.04.006> (2019).
10. Rothbart, M. K., Sheese, B. E., Rueda, M. R. & Posner, M. I. Developing mechanisms of self-regulation in early life. *Emot. Rev.* **3**, 207–213. <https://doi.org/10.1177/1754073910387943> (2011).
11. Ruff, H. A. Components of attention during infants' manipulative exploration. *Child Dev.* <https://doi.org/10.2307/1130642> (1986).
12. Swingler, M. M., Perry, N. B. & Calkins, S. D. Neural plasticity and the development of attention: Intrinsic and extrinsic influences. *Dev. Psychopathol.* **27**, 443–457. <https://doi.org/10.1017/s0954579415000085> (2015).
13. Nelson, C. A. & deRegnier, R. A. Neural correlates of attention and memory in the first year of life. *Dev. Neuropsychol.* **8**, 119–134. <https://doi.org/10.1080/87565649209540521> (1992).
14. Richards, J. E. Attention affects the recognition of briefly presented visual stimuli in infants: An ERP study. *Dev. Sci.* **6**, 312–328. <https://doi.org/10.1111/1467-7687.00287> (2003).
15. Mundy, P. & Jarrold, W. Infant joint attention, neural networks and social cognition. *Neural Netw.* **23**, 985–997. <https://doi.org/10.1016/j.neunet.2010.08.009> (2010).
16. Salley, B. et al. Infants' early visual attention and social engagement as developmental precursors to joint attention. *Dev. Psychol.* **52**, 1721–1731. <https://doi.org/10.1037/dev0000205> (2016).
17. Bradshaw, J., Fu, X. & Richards, J. E. Infant sustained attention differs by context and social content in the first 2 years of life. *Dev. Sci.* <https://doi.org/10.1111/desc.13500> (2024).
18. Dahl, A. Ecological commitments: Why developmental science needs naturalistic methods. *Child Dev. Perspect.* **11**, 79–84. <https://doi.org/10.1111/cdev.12217> (2016).
19. Casillas, M., Brown, P. & Levinson, S. C. Early language experience in a Tzeltal Mayan village. *Child Dev.* **91**, 1819–1835. <https://doi.org/10.1111/cdev.13349> (2019).
20. Casillas, M. & Elliott, M. Cross-cultural differences in children's object handling at home. *PsyArXiv* (2021). <https://doi.org/10.31234/osf.io/43db8>
21. Karasik, L. B., Tamis-LeMonda, C. S., Ossmy, O. & Adolph, K. E. The ties that bind: Cradling in Tajikistan. *PLoS ONE* <https://doi.org/10.1371/journal.pone.0204428> (2018).
22. Schroer, S. E., Peters, R. E. & Yu, C. Consistency and variability in multimodal parent–child social interaction: An at-home study using head-mounted eye trackers. *Dev. Psychol.* **60**, 1432 (2024).
23. McQuillan, M. E., Smith, L. B., Yu, C. & Bates, J. E. Parents influence the visual learning environment through children's manual actions. *Child Dev.* <https://doi.org/10.1111/cdev.13274> (2019).
24. Needham, A. Improvements in object exploration skills may facilitate the development of object segregation in early infancy. *J. Cogn. Dev.* **1**, 131–156. <https://doi.org/10.1207/s15327647jcd010201> (2000).
25. Suanda, S. H., Barnhart, M., Smith, L. B. & Yu, C. The signal in the noise: The visual ecology of parents' object naming. *Infancy* **24**, 455–476. <https://doi.org/10.1111/infa.12278> (2018).
26. Geangu, E. et al. EgoActive: Integrated wireless wearable sensors for capturing infant egocentric auditory–visual statistics and autonomic nervous system function 'in the wild'. *Sensors* <https://doi.org/10.3390/s23187930> (2023).
27. Zhu, Z., Liu, T., Li, G., Li, T. & Inoue, Y. Wearable sensor systems for infants. *Sensors* **15**, 3721–3749. <https://doi.org/10.3390/s150203721> (2015).
28. Long, B. L., Kachergis, G., Agrawal, K. & Frank, M. C. A longitudinal analysis of the social information in infants' naturalistic visual experience using automated detections. *Dev. Psychol.* **58**, 2211–2229. <https://doi.org/10.1037/dev0001414> (2022).
29. Kadooka, K. & Franchak, J. M. Developmental changes in infants' and children's attention to faces and salient regions vary across and within video stimuli. *Dev. Psychol.* **56**, 2065–2079. <https://doi.org/10.1037/dev0001073> (2020).
30. Papageorgiou, K. A. et al. Individual differences in infant fixation duration relate to attention and behavioral control in childhood. *Psychol. Sci.* **25**, 1371–1379. <https://doi.org/10.1177/0956797614531295> (2014).
31. Aneja, P. et al. Leveraging technological advances to assess dyadic visual cognition during infancy in high- and low-resource settings. *Front. Psychol.* <https://doi.org/10.3389/fpsyg.2024.1376552> (2024).
32. Erel, Y., Potter, C. E., Jaffe-Dax, S., Lew-Williams, C. & Bermato, A. H. iCatcher: A neural network approach for automated coding of young children's eye movements. *Infancy* **27**, 765–779. <https://doi.org/10.1111/infa.12468> (2022).
33. Colombo, J. On the neural mechanisms underlying developmental and individual differences in visual fixation in infancy: Two hypotheses. *Dev. Rev.* **15**, 97–135. <https://doi.org/10.1006/drev.1995.1005> (1995).
34. Richards, J. E. & Turner, E. D. Extended visual fixation and distractibility in children from six to twenty-four months of age. *Child Dev.* **72**, 963–972. <https://doi.org/10.1111/1467-8624.00328> (2001).
35. Richards, J. E. & Casey, B. J. Heart rate variability during attention phases in young infants. *Psychophysiology* **28**, 43–53 (1991).
36. Friedman, A. H., Watamura, S. E. & Robertson, S. S. Movement-attention coupling in infancy and attention problems in childhood. *Dev. Med. Child Neurol.* **47**, 660–665. <https://doi.org/10.1111/j.1469-8749.2005.tb01050.x> (2005).
37. Islam, B. et al. Preliminary technical validation of LittleBeats™: A multimodal sensing platform to capture cardiac physiology, motion, and vocalizations. *Sensors* <https://doi.org/10.3390/s24030901> (2024).
38. Maitha, C. et al. An open-source, wireless vest for measuring autonomic function in infants. *Behav. Res. Methods* **52**, 2324–2337. <https://doi.org/10.3758/s13428-020-01394-4> (2020).
39. Aminikhanghahi, S. & Cook, D. J. A survey of methods for time series change point detection. *Knowl. Inf. Syst.* **51**, 339–367. <https://doi.org/10.1007/s10115-016-0987-z> (2017).
40. Fryzlewicz, P. Detecting possibly frequent change-points: Wild Binary Segmentation 2 and steepest-drop model selection. *J. Korean Stat. Soci.* **49**, 1027–1070. <https://doi.org/10.1007/s42952-020-00060-x> (2020).
41. Slone, L. K. et al. Gaze in action: Head-mounted eye tracking of children's dynamic visual attention during naturalistic behavior. *JoVE* <https://doi.org/10.3791/58496> (2018).
42. Yu, C. & Smith, L. B. The social origins of sustained attention in one-year-old human infants. *Curr. Biol.* **26**, 1235–1240. <https://doi.org/10.1016/j.cub.2016.03.026> (2016).
43. Forte, G., Favieri, F. & Casagrande, M. Heart rate variability and cognitive function: A systematic review. *Front. Neurosci.* <https://doi.org/10.3389/fnins.2019.00710> (2019).
44. Thayer, J. F., Hansen, A. L., Saus-Rose, E. & Johnsen, B. H. Heart rate variability, prefrontal neural function, and cognitive performance: The neurovisceral integration perspective on self-regulation, adaptation, and health. *Ann. Behav. Med.* **37**, 141–153. <https://doi.org/10.1007/s12160-009-9101-z> (2009).
45. Akselrod, S. et al. Power spectrum analysis of heart rate fluctuation: A quantitative probe of beat-to-beat cardiovascular control. *Science* **213**, 220–222. <https://doi.org/10.1126/science.6166045> (1981).
46. Hansen, A. L., Johnsen, B. H. & Thayer, J. F. Vagal influence on working memory and attention. *Int. J. Psychophysiol.* **48**, 263–274. [https://doi.org/10.1016/s0167-8760\(03\)00073-4](https://doi.org/10.1016/s0167-8760(03)00073-4) (2003).
47. Börger, N. et al. Heart rate variability and sustained attention in ADHD children. *J. Abnorm. Child Psychol.* **27**, 25–33 (1999).
48. Van Roon, A. M., Mulder, L. J. M., Althaus, M. & Mulder, G. Introducing a baroreflex model for studying cardiovascular effects of mental workload. *Psychophysiology* **41**, 961–981. <https://doi.org/10.1111/j.1469-8986.2004.00251.x> (2004).
49. Griffiths, K. R. et al. Sustained attention and heart rate variability in children and adolescents with ADHD. *Biol. Psychol.* **124**, 11–20. <https://doi.org/10.1016/j.biopsycho.2017.01.004> (2017).

50. Addison, P. S. Wavelet transforms and the ECG: A review. *Physiol. Meas.* **26**, R155–R199. <https://doi.org/10.1088/0967-3334/26/5/r01> (2005).
51. Pichot, V. et al. Wavelet transform to quantify heart rate variability and to assess its instantaneous changes. *J. Appl. Physiol.* **86**, 1081–1091. <https://doi.org/10.1152/jappl.1999.86.3.1081> (1999).
52. Toledo, E., Gurevitz, O., Hod, H., Eldar, M. & Akselrod, S. Wavelet analysis of instantaneous heart rate: A study of autonomic control during thrombolysis. *Am. J. Physiol. -Regulat., Integrat. Comparat. Physiol.* **284**, R1079–R1091. <https://doi.org/10.1152/ajpregu.00287.2002> (2003).
53. Zhang, Y. S., Takahashi, D. Y., El Hady, A., Liao, D. A. & Ghazanfar, A. A. Active neural coordination of motor behaviors with internal states. *Proceed. Nat. Acad. Sci.* <https://doi.org/10.1073/pnas.2201194119> (2022).
54. Amso, D., Haas, S. & Markant, J. An eye tracking investigation of developmental change in bottom-up attention orienting to faces in cluttered natural scenes. *PLoS ONE* **9**, e85701 (2014).
55. Byrge, L., Sporns, O. & Smith, L. B. Developmental process emerges from extended brain–body–behavior networks. *Trends Cogn. Sci.* **18**, 395–403 (2014).
56. Crespo-Llado, M. M., Vanderwert, R., Roberti, E. & Geangu, E. Eight-month-old infants' behavioral responses to peers' emotions as related to the asymmetric frontal cortex activity. *Sci. Report* <https://doi.org/10.1038/s41598-018-35219-4> (2018).
57. DeBolt, M. C. et al. A new perspective on the role of physical salience in visual search: Graded effect of salience on infants' attention. *Dev. Psychol.* **59**, 326–343. <https://doi.org/10.1037/dev0001460> (2023).
58. Geangu, E. & Vuong, Q. C. Look up to the body: An eye-tracking investigation of 7-months-old infants' visual exploration of emotional body expressions. *Infant Behav. Dev.* <https://doi.org/10.1016/j.infbeh.2020.101473> (2020).
59. Geangu, E. & Vuong, Q. C. Seven-months-old infants show increased arousal to static emotion body expressions: Evidence from pupil dilation. *Infancy* **28**, 820–835. <https://doi.org/10.1111/inf.12535> (2023).
60. Kwon, M.-K., Setoodehnia, M., Baek, J., Luck, S. J. & Oakes, L. M. The development of visual search in infancy: Attention to faces versus salience. *Dev. Psychol.* **52**, 537–555. <https://doi.org/10.1037/dev0000080> (2016).
61. Oakes, L. M., Hayes, T. R., Klotz, S. M., Pomaranski, K. I. & Henderson, J. M. The role of local meaning in infants' fixations of natural scenes. *Infancy* **29**, 284–298. <https://doi.org/10.1111/inf.12582> (2024).
62. Sun, L., Francis, D. J., Nagai, Y. & Yoshida, H. Early development of saliency-driven attention through object manipulation. *Acta Psychol. (Oxf)* <https://doi.org/10.1016/j.actpsy.2024.104124> (2024).
63. Tummelshammer, K. S., Wu, R., Sobel, D. M. & Kirkham, N. Z. Infants track the reliability of potential informants. *Psychol. Sci.* **25**, 1730–1738. <https://doi.org/10.1177/0956797614540178> (2014).
64. van Renswoude, D. R., Visser, I., Raijmakers, M. E. J., Tsang, T. & Johnson, S. P. Real-world scene perception in infants: What factors guide attention allocation? *Infancy* **24**, 693–717. <https://doi.org/10.1111/inf.12308> (2019).
65. Harel, J., Koch, C. & Perona, P. Graph-based visual saliency. *Advances in Neural Information Processing Systems* **19** (2006).
66. Rosenholtz, R., Li, Y. & Nakano, L. Measuring visual clutter. *J. Vision* <https://doi.org/10.1167/7.2.17> (2007).
67. Mason, H. T. et al. A complete pipeline for heart rate extraction from infant ECG. (2024).
68. Borjon, J. I., Abney, D. H., Yu, C. & Smith, L. B. Head and eyes: Looking behavior in 12- to 24-month-old infants. *J. Vision* <https://doi.org/10.1167/jov.21.8.18> (2021).
69. Yu, C. & Smith, L. B. Joint attention without gaze following: Human infants and their parents coordinate visual attention to objects through eye-hand coordination. *PLoS ONE* **8**, e79659 (2013).
70. Wass, S. V., Clackson, K. & de Barbaro, K. Temporal dynamics of arousal and attention in 12-month-old infants. *Dev. Psychobiol.* **58**, 623–639. <https://doi.org/10.1002/dev.21406> (2016).
71. Wass, S. V., de Barbaro, K. & Clackson, K. Tonic and phasic co-variation of peripheral arousal indices in infants. *Biol. Psychol.* **111**, 26–39. <https://doi.org/10.1016/j.biopsycho.2015.08.006> (2015).
72. Makowski, D. et al. NeuroKit2: A Python toolbox for neurophysiological signal processing. *Behav. Res. Methods* **53**, 1689–1696. <https://doi.org/10.3758/s13428-020-01516-y> (2021).
73. Lansink, J. M. & Richards, J. E. Heart rate and behavioral measures of attention in six-, nine-, and twelve-month-old infants during object exploration. *Child Dev.* **68**, 610–620. <https://doi.org/10.2307/1132113> (1997).
74. Lansink, J. M., Mintz, S. & Richards, J. E. The distribution of infant attention during object examination. *Dev. Sci.* **3**, 163–170. <https://doi.org/10.1111/1467-7687.00109> (2000).
75. Richards, J. E. Localization of peripheral stimuli by infants: Age, attention and individual differences in heart rate variability. *J. Exp. Psychol. Hum. Percept. Perform.* **23**, 667–680 (1997).
76. Richards, J. E. Infant visual sustained attention and respiratory sinus arrhythmia. *Child Dev.* <https://doi.org/10.2307/1130525> (1987).
77. Richards, J. E. Development of selective attention in young infants: Enhancement and attenuation of startle reflex by attention. *Dev. Sci.* **1**, 45–51 (1998).
78. Petrie Thomas, J. H., Whitfield, M. F., Oberlander, T. F., Synnes, A. R. & Grunau, R. E. Focused attention, heart rate deceleration, and cognitive development in preterm and full-term infants. *Dev. Psychobiol.* **54**, 383–400. <https://doi.org/10.1002/dev.20597> (2012).
79. Anderson, E. M., Seemiller, E. S. & Smith, L. B. Scene saliencies in egocentric vision and their creation by parents and infants. *Cognition* <https://doi.org/10.1016/j.cognition.2022.1052563> (2022).
80. Sun, L. & Yoshida, H. Why the parent's gaze is so powerful in organizing the infant's gaze: The relationship between parental referential cues and infant object looking. *Infancy* **27**, 780–808. <https://doi.org/10.1111/inf.12475> (2022).
81. Pérez-Edgar, K. et al. Patterns of sustained attention in infancy shape the developmental trajectory of social behavior from toddlerhood through adolescence. *Dev. Psychol.* **46**, 1723–1730. <https://doi.org/10.1037/a0021064> (2010).
82. Richards, J. E. The development of attention to simple and complex visual stimuli in infants: Behavioral and psychophysiological measures. *Dev. Rev.* **30**, 203–219. <https://doi.org/10.1016/j.dr.2010.03.005> (2010).
83. Richards, J. E. & Cronise, K. Extended visual fixation in the early preschool years: Look duration, heart rate changes, and attentional inertia. *Child Dev.* **71**, 602–620. <https://doi.org/10.1111/1467-8624.00170> (2000).
84. Richards, J. E. & Gibson, T. L. Extended visual fixation in young infants: Look distributions, heart rate changes, and attention. *Child Dev.* <https://doi.org/10.2307/1132290> (1997).
85. Luo, C. & Franchak, J. M. Head and body structure infants' visual experiences during mobile, naturalistic play. *PLoS ONE* **15**, e0242009 (2020).
86. Hunter, B. K., Klotz, S., DeBolt, M., Luck, S. & Oakes, L. Evaluation of graph-based visual saliency model using infant fixation data. *J. Vision* <https://doi.org/10.1167/jov.23.9.5704> (2023).
87. Itti, L. & Koch, C. Computational modelling of visual attention. *Nat. Rev. Neurosci.* **2**, 194–203. <https://doi.org/10.1038/35058500> (2001).
88. Mahdi, A., Su, M., Schlesinger, M. & Qin, J. A comparison study of saliency models for fixation prediction on infants and adults. *IEEE Trans. Cognit. Dev. Syst.* **10**, 485–498. <https://doi.org/10.1109/tcds.2017.2696439> (2018).
89. Franchak, J. M., Kretch, K. S., Soska, K. C. & Adolph, K. E. Head-mounted eye tracking: A new method to describe infant looking. *Child Dev.* **82**, 1738–1750. <https://doi.org/10.1111/j.1467-8624.2011.01670.x> (2011).
90. Kretch, K. S. & Adolph, K. E. The organization of exploratory behaviors in infant locomotor planning. *Dev. Sci.* <https://doi.org/10.1111/desc.12421> (2017).

91. Park, T., Eckley, I. A. & Ombao, H. C. Dynamic classification using multivariate locally stationary wavelet processes. *Signal Process.* **152**, 118–129. <https://doi.org/10.1016/j.sigpro.2018.01.005> (2018).
92. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B Stat Methodol.* **58**, 267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x> (1996).
93. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **16**, 321–357. <https://doi.org/10.1613/jair.953> (2002).
94. Smith, L. B., Jayaraman, S., Clerkin, E. & Yu, C. The developing infant creates a curriculum for statistical learning. *Trends Cogn. Sci.* **22**, 325–336. <https://doi.org/10.1016/j.tics.2018.02.004> (2018).
95. Smith, L. B. & Karmazyn-Raz, H. Episodes of experience and generative intelligence. *Trends Cogn. Sci.* **26**, 1064–1065 (2022).
96. Sparks, R. Z. et al. in *Proceedings of the Annual Meeting of the Cognitive Science Society*.
97. Sullivan, J., Mei, M., Perfors, A., Wojcik, E. & Frank, M. C. SAYCam: A large, longitudinal audiovisual dataset recorded from the infant's perspective. *Open Mind* **5**, 20–29. https://doi.org/10.1162/opmi_a_00039 (2021).
98. Aci, Ç. İ., Kaya, M. & Mishchenko, Y. Distinguishing mental attention states of humans via an EEG-based passive BCI using machine learning methods. *Expert Syst. Appl.* **134**, 153–166. <https://doi.org/10.1016/j.eswa.2019.05.057> (2019).
99. Belle, A., Hargraves, R. H. & Najarian, K. An automated optimal engagement and attention detection system using electrocardiogram. *Comput. Math. Methods Med.* **2012**, 1–12. <https://doi.org/10.1155/2012/528781> (2012).
100. Chen, C. M., Wang, J. Y. & Yu, C. M. Assessing the attention levels of students by using a novel attention aware system based on brainwave signals. *Br. J. Edu. Technol.* **48**, 348–369. <https://doi.org/10.1111/bjet.12359> (2015).
101. Khare, S. K., Bajaj, V., Gaikwad, N. B. & Sinha, G. R. Ensemble wavelet decomposition-based detection of mental states using electroencephalography signals. *Sensors* <https://doi.org/10.3390/s23187860> (2023).
102. Liu, N.-H., Chiang, C.-Y. & Chu, H.-C. Recognizing the degree of human attention using EEG signals from mobile sensors. *Sensors* **13**, 10273–10286. <https://doi.org/10.3390/s130810273> (2013).
103. Koh, J. E. W. et al. Automated classification of attention deficit hyperactivity disorder and conduct disorder using entropy features with ECG signals. *Comput. Biol. Med.* <https://doi.org/10.1016/j.combiomed.2021.105120> (2022).
104. Tilwani, D., Bradshaw, J., Sheth, A. & O'Reilly, C. ECG recordings as predictors of very early autism likelihood: a machine learning approach. *Bioengineering* <https://doi.org/10.3390/bioengineering10070827> (2023).
105. Borjon, J. I., Takahashi, D. Y., Cervantes, D. C. & Ghazanfar, A. A. Arousal dynamics drive vocal production in marmoset monkeys. *J. Neurophysiol.* **116**, 753–764. <https://doi.org/10.1152/jn.00136.2016> (2016).
106. Finley, J. P. & Nugent, S. T. Heart rate variability in infants, children and young adults. *J. Auton. Nerv. Syst.* **51**, 103–108. [https://doi.org/10.1016/0165-1838\(94\)00117-3](https://doi.org/10.1016/0165-1838(94)00117-3) (1995).
107. Byrne, J. M. & Smith-Martel, D. J. Cardiac-somatic integration: An index of visual attention. *Infant Behav. Dev.* **10**, 493–500 (1987).
108. Chen, Y., Matheson, L. E. & Sakata, J. T. Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proc. Natl. Acad. Sci.* **113**, 6641–6646. <https://doi.org/10.1073/pnas.1522306113> (2016).
109. Bazhenova, O. V., Plonskaia, O. & Porges, S. W. Vagal reactivity and affective adjustment in infants during interaction challenges. *Child Dev.* **72**, 1314–1326. <https://doi.org/10.1111/1467-8624.00350> (2003).
110. Porges, S. W. et al. Does motor activity during psychophysiological paradigms confound the quantification and interpretation of heart rate and heart rate variability measures in young children?. *Dev. Psychobiol.* **49**, 485–494. <https://doi.org/10.1002/dev.20228> (2007).
111. Gutierrez-Barragan, D., Basson, M. A., Panzeri, S. & Gozzi, A. Infraslow state fluctuations govern spontaneous fMRI network dynamics. *Curr. Biol.* **29**, 2295–2306 (2019).
112. Raut, R. V. et al. Global waves synchronize the brain's functional systems with fluctuating arousal. *Sci. Adv.* <https://doi.org/10.1126/sciadv.abf2709> (2021).
113. Pomaranski, K. I., Hayes, T. R., Kwon, M.-K., Henderson, J. M. & Oakes, L. M. Developmental changes in natural scene viewing in infancy. *Dev. Psychol.* **57**, 1025 (2021).
114. Helo, A., van Ommen, S., Pannasch, S., Danten-Dordoigne, L. & Rämä, P. Influence of semantic consistency and perceptual features on visual attention during scene viewing in toddlers. *Infant Behav. Dev.* **49**, 248–266. <https://doi.org/10.1016/j.infbeh.2017.09.008> (2017).
115. Ruff, H. A. & Rothbart, M. K. *Attention in early development: Themes and variations* (Oxford University Press, 2001).
116. Adolph, K. E., Cole, W. G. & Vereijken, B. Intraindividual variability in the development of motor skills in childhood. *Handbook of intraindividual variability across the life span*, 59–83 (2014).
117. Atun-Einy, O., Berger, S. E. & Scher, A. Pulling to stand: Common trajectories and individual differences in development. *Dev. Psychobiol.* **54**, 187–198. <https://doi.org/10.1002/dev.20593> (2011).
118. Nakagawa, A. et al. Investigating the link between temperamental and motor development: a longitudinal study of infants aged 6–42 months. *BMC Pediatr.* <https://doi.org/10.1186/s12887-024-05038-w> (2024).
119. Graham, F. K. & Jackson, J. C. *Advances in Child Development and Behavior* Vol. 5, 59–117 (Elsevier, 1970).
120. Borjon, J. I., Sahoo, M. K., Rhodes, K. D., Lipschutz, R. & Bick, J. R. Recognizability and timing of infant vocalizations relate to fluctuations in heart rate. *Proceed. Nat. Acad. Sci.* <https://doi.org/10.1073/pnas.2419650121> (2024).
121. Gartstein, M. A. & Rothbart, M. K. Studying infant temperament via the revised infant behavior questionnaire. *Infant Behav. Dev.* **26**, 64–86. [https://doi.org/10.1016/s0163-6383\(02\)00169-8](https://doi.org/10.1016/s0163-6383(02)00169-8) (2003).
122. Peltola, M. J., Leppänen, J. M. & Hietanen, J. K. Enhanced cardiac and attentional responding to fearful faces in 7-month-old infants. *Psychophysiology* **48**, 1291–1298. <https://doi.org/10.1111/j.1469-8986.2011.01188.x> (2011).

Acknowledgements

We would like to express our gratitude to all families who dedicated their time to participate in this research. Without their continued interest in our research and desire to help, these findings would have not been possible. The authors would also like to thank Brigita Ceponyte, Emily Clayton, Nicoleta Gavrila, Aastha Mishra, Lauren Charters, Fiona Frame, and Laura Tissiman for their invaluable support and help throughout the project.

Author contributions

Conceptualization: EG, PMC, YZ, HM, MIK, QCV. Methodology: EG, PMC, YZ, QCV, HM, MIK, MCG, DM. Formal Analysis: YZ, QCV, HM, MIK, PMC, EG. Investigation: EG, PMC, MCG. Writing – Original Draft Preparation: EG, PMC, YZ, HM, MIK, QCV. Software: YZ, HM, QCV, DM. Validation: EG, PMC, YZ, HM, QCV. Resources: EG, QCV, MIK. Data Curation: EG, QCV, HM, PMC. Visualisation: YZ, QCV, HM, EG. Supervision: EG, QCV, MIK. Project Administration: EG. Funding Acquisition: EG, QCV, MIK.

Funding

The work presented in this paper was partially funded by Wellcome Leap - The 1kD Program.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-96794-x>.

Correspondence and requests for materials should be addressed to E.G.

Reprints and permissions information is available at www.nature.com/reprints.

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025